

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2005, 21:37:29 ; Search time 52.7808 Seconds  
(without alignments)  
291.060 Million cell updates/sec

Title: 09782816-1-22

Perfect score: 138

Sequence: 1 EVEKIKTTVKESATREKLTPTLAKOLAAI 30

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: UniProt\_03:\*

1: UniProt\_sprot:\*

2: UniProt\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	138	100.0	400	1	DCT2_HUMAN
2	135	97.8	401	1	DCT2_MOUSE
3	135	97.8	402	1	DCT2_MOUSE
4	114	82.6	338	2	Q6AYH5
5	114	82.6	403	2	Q7ZXY2
6	109	79.0	403	2	Q6IRB3
7	92	66.7	403	2	Q6GJ30
8	89	64.5	402	2	Q7T3H1
9	56	40.6	333	2	Q9PTG6
10	54	39.1	445	2	Q874E7
11	54	39.1	500	1	C7ZY ARATH
12	54	39.1	769	2	Q6BPD3
13	54	39.1	933	1	GLND_AZORR
14	53	38.4	416	2	Q6BFZ2
15	53	38.4	710	2	Q6CFJ2
16	53	38.4	757	2	Q7SCZ5
17	53	38.4	1373	2	Q01787
18	52	37.7	449	2	Q738U0
19	52	37.7	452	2	Q81DX2
20	52	37.7	470	2	Q6FAP3
21	52	37.7	1430	2	Q7RT55
22	51	37.0	131	2	Q8T2Z5
23	51	37.0	180	2	Q2S636
24	51	37.0	319	2	Q8PST8
25	51	37.0	378	2	Q6UYC7
26	51	37.0	389	2	Q8KH88
27	51	37.0	389	2	Q8LON4
28	51	37.0	389	2	Q8LON6
29	51	37.0	389	2	Q8LON9
30	51	37.0	389	2	Q8LON0
31	51	37.0	389	2	Q8LON1

X=0  
K=1  
L=1  
M=1  
N=1  
O=1  
P=1  
Q=1  
R=1  
S=1  
T=1  
U=1  
V=1  
W=1  
X=1  
Y=1  
Z=1

32	51	37.0	389	2	Q8L0P2
33	51	37.0	389	2	Q8L0P5
34	51	37.0	389	2	Q8L0P7
35	51	37.0	500	1	C7ZV ARATH
36	51	37.0	757	2	Q6GWT0
37	51	37.0	770	2	Q7ZFA2
38	51	37.0	773	1	Q7RB_HELPJ
39	51	37.0	928	2	Q6OE1
40	51	37.0	932	2	Q51674
41	51	37.0	964	1	IP2_CHRYO
42	51	37.0	10917	2	Q93NM6
43	50.5	36.6	213	1	LIPB_SALTI
44	50.5	36.6	213	1	LIPB_SALTY
45	50.5	36.6	380	2	Q7MPJ5

## ALIGNMENTS

RESULT 1

DCT2\_HUMAN

ID DCT2\_HUMAN STANDARD; PRT; 400 AA.

AC Q13561; Q86YN2; Q9BW17;

DT 01-NOV-1997 (Rel. 35, Created)

DT 05-JUL-2004 (Rel. 44, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Dynactin complex 50 kDa subunit (50 kDa dynein-associated polypeptide)

DE (p50 dynactin) (DCTN-50) (Dynactin 2).

GN Name=DCTN2; Synonyms=DCTN50;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OC NCBI TaxID=9606;

OX [1]

RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.

RX MEDLINE=96178072; PubMed=8647893; DOI=10.1083/jcb.132.4.617;

RT Echeverri C.J., Paschal B.M., Vaughan K.T., Vallee R.B.;

RT "Molecular characterization of the 50-kD subunit of dynactin reveals function for the complex in chromosome alignment and spindle organization during mitosis."

RT J Cell Biol. 132:617-633(1996).

RL [2]

RN SEQUENCE FROM N.A.

RP TISSUE=Placenta, Skin, and Uterus;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Ustin T.B., Toshitsuki S., Carrinci P., Prange C.,

RA Rana S.S., Loguelfano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huij R.A.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,

RA Schercher A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [3]

RP SEQUENCE OF 1-384 FROM N.A.

RA Aumais J.P., Yu-Dee L.-Y.;

RT "Human 50 kD dynactin subunit, p50 dynactin, isolated from HeLa

RT cells."

RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE OF 1-13.

```

CC TISSUE=Platelet;
RX MEDLINE=22608298; PubMed=12665801; DOI=10.1038/nbt810;
RA Gvaert K., Goethals M., Martens L., Van Damme J., Stes A.,
RA Thomas G.R., Vandekerckhove J.;
RT "Exploring proteomes and analyzing protein processing by mass
RT spectrometric identification of sorted N-terminal peptides.";
RL Nat. Biotechnol. 21:566-569(2003).
CC -1- FUNCTION: Modulates cytoplasmic dynein binding to an organelle,
CC and plays a role in prometaphase chromosome alignment and spindle
CC organization during mitosis. May play a role in synapse formation
CC during brain development.
CC -1- SUBUNIT: Subunit of dyneactin, a multiprotein complex associated
CC with dynein.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; U50733; AAC0423.1; -.
CC EMBL; BC000718; AAH00718.1; -.
CC EMBL; BC009468; AAH09468.1; -.
CC EMBL; BC014083; AAH14083.1; -.
CC EMBL; AY189155; AAO34395.1; -.
CC Genbank; HGNC:2712; DCTN2.
CC
CC MIM; 607376; -.
CC DR GO; GO:0005813; C:centrosome; TAS.
CC DR GO; GO:0005869; C:dyneactin complex; TAS.
CC DR GO; GO:0000776; C:kinechochore; TAS.
CC DR GO; GO:0008283; P:cell proliferation; TAS.
CC DR GO; GO:0007067; P:mitosis; TAS.
CC DR InterPro; IPR006996; Dyneactin.
CC DR Pfam; PF04912; Dyneactin; 1.
CC DR Coiled coil; Cytokeleton; Direct protein sequencing; Dynein;
CC Membrane; Microtubule; Motor protein.
CC
CC FT INIT MET 0
CC FT DOMAIN 98 131 Coiled coil (Potential).
CC FT DOMAIN 213 243 Coiled coil (Potential).
CC FT DOMAIN 378 398 Coiled coil (Potential).
CC FT CONFLICT 34 34 A -> AFADEL (in Ref. 1).
CC FT CONFLICT 35 35 E -> ELE (in Ref. 3).
CC FT CONFLICT 381 384 LATV -> PGHS (in Ref. 3).
CC SQ SEQUENCE 400 AA; 44099 MW; 0A95A893C0B8270F CRC64;
CC
CC Query Match 100.0%; Score 138; DB 1; Length 400;
CC Best Local Similarity 100.0%; Pred. No. 1.8e-09;
CC Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC Oy 1 EVEKIKTVKESATSEKTPVILAKQALAL 30
CC | | | | | | | | | | | | | | | | | |
CC Db 115 EVEKIKTVKESATSEKTPVILAKQALAL 144
CC
CC RESULT 2
CC DCT2_MOUSE STANDARD; PRT; 401 AA.
CC AC Q99KJ8;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 05-JUL-2004 (Rel. 44, Last sequence update)
CC DT 05-JUL-2004 (Rel. 44, Last annotation update)
CC DE Dyneactin complex 50 kDa subunit (50 kDa dynein-associated polypeptide)
CC DE (p50 dyneactin) (DCTN-50) (Dyactin 2) (Growth cone membrane protein
CC DE 23-48K) (GMP23-48K).
CC GN Name=Dctn2;
CC OS Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC OC NCBI_TaxID=10090;
CC RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Cassavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shcherbino Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schercher A., Schein J.E., Jones S.J.M., Maira M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC
CC [2]
CC SEQUENCE OF 65-74; 77-91; 102-116; 156-170; 194-216 AND 309-320,
CC SUBCELLULAR LOCATION, AND DEVELOPMENTAL STAGE.
CC RP TISSUE=Brain;
CC RX MEDLINE=97289622; PubMed=9144527; DOI=10.1006/birc.1997.6447;
CC ABT T.K., Tanaka H., Iwanaga T., Odani S., Kuwano R.;
CC "The presence of the 50-kDa subunit of dyneactin complex in the nerve
CC growth cone.";
CC RL Biochem. Biophys. Res. Commun. 233:295-299(1997).
CC
CC [3]
CC INTERACTION WITH BICD2.
CC RP MEDLINE=21376052; PubMed=11483508; DOI=10.1093/emboj/20.15.4041;
CC Hoogenraad C.C., Akmanova A., Howell S.A., Dordland B.R.,
CC de Zeeuw C.I., Willemsen R., Visser P., Grosfeld F., Galjart N.;
CC "Mammalian Golgi-associated Bicaudal-D2 functions in the dynein-
CC dyneactin pathway by interacting with these complexes.";
CC RL EMBO J. 20:4041-4054(2001).
CC
CC -1- FUNCTION: Modulates cytoplasmic dynein binding to an organelle,
CC and plays a role in prometaphase chromosome alignment and spindle
CC organization during mitosis. May play a role in synapse formation
CC during brain development.
CC -1- SUBUNIT: Subunit of dyneactin, a multiprotein complex associated
CC with dynein (By similarity). Interacts with BICD2.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
CC -1- DEVELOPMENTAL STAGE: Present at high levels in both cytoplasmic
CC and membrane-associated forms in neonates. Levels of membrane-
CC associated form are greatly reduced in the adult.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; BC004613; AAH04613.1; -.
CC DR MGD; MGI:107733; Dctn2.
CC DR InterPro; IPR006996; Dyneactin.
CC DR Pfam; PF04912; Dyneactin; 1.
CC DR Coiled coil; Cytokeleton; Direct protein sequencing; Dynein;
CC Membrane; Microtubule; Motor protein.
CC FT INIT MET 0
CC FT DOMAIN 98 131 By similarity. Coiled coil (Potential).
CC FT DOMAIN 214 244 Coiled coil (Potential).
CC SQ SEQUENCE 401 AA; 43985 MW; 1535E4ABD5940BEC CRC64;
CC
CC Query Match 97.8%; Score 135; DB 1; Length 401;
CC Best Local Similarity 96.7%; Pred. No. 4.5e-09;
CC Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 EVEKITTVESATEEKLTPVLA KOALA 30  
 DB 115 EVEKITTVESATEEKLTPVLA KOALA 144  
 RESULT 3  
 ID 06AYHS PRELIMINARY; PRT; 402 AA.  
 AC 06AYHS  
 DT 25-OCT-2004 (TREMBlrel. 28, Created)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DE Dynactin 2.  
 GN Name=Dctn2;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NC NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Abramson R.D., Mullahy S.J.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,  
 RA Krywinski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RC Director AGC Project;  
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC079042; AAH79042.1;  
 DR GO: GO:0005869; C:dynactin complex; IEA.  
 DR GO: GO:0007017; P:microtubule-based process; IEA.  
 DR InterPro: IPR006996; Dynactin.  
 DR Pfam: PF04912; Dynactin; 1.  
 SQ SEQUENCE 402 AA; 44148 MW; 550335535A4FB052 CRC64;  
 Query Match 97.8%; Score 135; DB 2; Length 402;  
 Best Local Similarity 96.7%; Pred. No. 4.5e-09;  
 Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EVEKITTVESATEEKLTPVLA KOALA 30  
 DB 116 EVEKITTVESATEEKLTPVLA KOALA 145  
 RESULT 4  
 ID 07ZXY2 PRELIMINARY; PRT; 338 AA.  
 AC 07ZXY2  
 DT 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Dctn2-p10v protein.  
 OS Xenopus laevis (African clawed frog).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 NC NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Abramson R.D., Mullahy S.J.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,  
 RA Krywinski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S.D., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative.";  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RC Klein S., Strausberg R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC044069; AAH44069.1;  
 DR GO: GO:0005869; C:dynactin complex; IEA.  
 DR GO: GO:0007017; P:microtubule-based process; IEA.  
 DR InterPro: IPR006996; Dynactin.  
 DR Pfam: PF04912; Dynactin; 1.  
 SQ SEQUENCE 338 AA; 38257 MW; 334E0F0F644B186 CRC64;  
 Query Match 82.6%; Score 114; DB 2; Length 338;  
 Best Local Similarity 80.0%; Pred. No. 2e-06;  
 Matches 24; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 EVEKITTVESATEEKLTPVLA KOALA 30  
 DB 51 EVEKITTVESATEEKLTPVLA KOALA 80  
 RESULT 5  
 ID 06IRB3 PRELIMINARY; PRT; 403 AA.  
 AC 06IRB3  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Dctn2-p10v protein.  
 GN Name=dctn2-p10v;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 NC NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;  
RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hejich F.,  
RA Datchenko L., Marmura K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Cavaant T.L., Schetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Rata S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Ramey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman T.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E.,  
RA Jones S.J., Mair M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.,  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT initiative."  
RL Dev. Dyn. 225:384-391(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RL Klein S., Strausberg R.,  
RL Submitted (May-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL, BC070987, AA070987.1; -.  
DR GO; GO:0005865; C:dynactin complex; IEA.  
DR GO; GO:0007017; P:microtubule-based process; IEA.  
DR InterPro; IPR006996; Dynactin.  
DR Pfam; PF04912; Dynactin; 1.  
SQ SEQUENCE 403 AA; 44828 MW; 9FD158C7C7983062 CRC64;

Query Match 82.6%; Score 114; DB 2; Length 403;  
Best Local Similarity 80.0%; Pred. No. 2.3e-06;  
Matches 24; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVEKIKTVKESATEEKLTPVILAAQAL 30  
DB 116 EVEKIQSTLKESATEEKLTPVALAKQVAL 145

RESULT 6

O66J30  
ID O66J30 PRELIMINARY; PRT; 403 AA.  
AC O66J30;  
DT 25-OCT-2004 (TREMBLrel. 28, Created)  
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)  
DE MGCR2128 protein.  
GN Name=MGCR2128;  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;  
OC Xenopodina; Xenopus.  
OX NCBI\_taxid=8335;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.,  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT initiative."

```

RT      initiativc."
RL      Dev. Dyn. 225:384-391 (2002).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Kidney;
RX      PubMed=12477932; DOI=10.1073/pnas.242603899;
RA      Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuller G.D.,
RA      Altschul S.F., Zeeberg B., Butlow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Datchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA      Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA      Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
RA      Villallon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahy J., Helton W., Kettleman M., Madan A., Rodriguez S., Sanchez A.,
RA      Whitling M., Madan A., Young A.C., Shvachenko Y., Bouffard G.G.,
RA      Bakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Gilwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA      Rzywnicki M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA      Jones S.J., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length human
RT      and mouse cDNA sequences."
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Kidney;
RX      Klein S., Gerhard D.S.;
RL      Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR      EMBL, BC081081, AA081081.1;
DR      InterPro; IPR006996, Dymamltin.
DR      Pfam; PF04912; Dymamltin; 1.
SQ      SEQUENCE 403 AA; 44737 MW; 86BB80CB54325EF3 CRC64;

Query Match      79.0%; Score 109; DB 2; Length 403;
Best Local Similarity 76.7%; Pred. No. 1e-05;
Matches 23; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

CY      1 EVEKIKTTVKESATEEKLTPVLAKKQIAAL 30
Db      116 EKEKTIQSTVKSAAEKLTPVAAKQVNSL 145

RESULT 7
0773H1 PRELIMINARY; PRT; 405 AA.
AC      0773H1.
DT      01-OCT-2003 (TREMBlrel. 25, Created)
DT      01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT      01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE      Similar to dymactin 2 (P50).
GN      ORRnamsa-zgc:63867,
OS      Brachydanio rerio (Zebrafish) (Danio rerio).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Osteichthys; Cypriniformes;
OC      Cyprinidae; Danio.
OX      NCBI_TaxID=7955;
[1]
RN      RP      SEQUENCE FROM N.A.
RC      TISSUE=Embryo;
RX      MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA      Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuller G.D.,
RA      Altschul S.F., Zeeberg B., Butlow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Datchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA      Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA      Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
RA      Villallon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahy J., Helton W., Kettleman M., Madan A., Rodriguez S., Sanchez A.,
RA      Whitling M., Madan A., Young A.C., Shvachenko Y., Bouffard G.G.,
RA      Bakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Gilwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA      Rzywnicki M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA      Jones S.J., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length human
RT      and mouse cDNA sequences."
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Kidney;
RX      Klein S., Gerhard D.S.;
RL      Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR      EMBL, BC081081, AA081081.1;
DR      InterPro; IPR006996, Dymamltin.
DR      Pfam; PF04912; Dymamltin; 1.
SQ      SEQUENCE 403 AA; 44737 MW; 86BB80CB54325EF3 CRC64;

```



RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywnski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Maira M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Straussberg R.;  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC053120; AAH53120.1; -  
 DR ZFIN; ZDB-GENE-040426-1279; zgc:63867.  
 DR GO; GO:0005869; C:dynactin complex; IEA.  
 DR GO; GO:0007017; P:microtubule-based process; IEA.  
 DR InterPro; IPR006996; Dynactin.  
 DR Pfam; Pf04912; Dynactin; 1.  
 SQ SEQUENCE 405 AA; 44625 MW; CA00047342500953 CRC64;

Query Match 66.7%; Score 92; DB 2; Length 405;  
 Best Local Similarity 63.3%; Pred. No. 0.0016;  
 Matches 19; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 1 EVEKIKTVKESATEEKLTPVLAQKQAL 30  
 Db 118 DVEKIQSTKESAEERLTPVLAQQAQL 147

## RESULT 8

Q9PTG6 PRELIMINARY; PRT; 402 AA.  
 AC Q9PTG6;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Dynactin.  
 GN Name=D50;  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianidae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20056093; PubMed=10588646;  
 RA Valetti C., Weizel D.M., Schrader M., Hasbani M.J., Gill S.R.,  
 RA Kreis T.B., Schroer T.A.;  
 RT "Role of dynactin in endocytic traffic: effects of dynactin  
 RT overexpression and colocalization with CLIP-170.";  
 RL Mol. Biol. Cell 10:4107-4120 (1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Schroer T.A., Gill S.R., Hasbani J., Crego C.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF200744; AAF1396.1; -  
 DR GO; GO:0005869; C:dynactin complex; IEA.  
 DR GO; GO:0007017; P:microtubule-based process; IEA.  
 DR InterPro; IPR006996; Dynactin.  
 DR Pfam; Pf04912; Dynactin; 1.  
 SQ SEQUENCE 402 AA; 45126 MW; F2296467C630DCB9 CRC64;

Query Match 64.5%; Score 89; DB 2; Length 402;  
 Best Local Similarity 60.0%; Pred. No. 0.0039;  
 Matches 18; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EVEKIKTVKESATEEKLTPVLAQKQAL 30  
 Db 118 DVEKIQSAVKESAEERLTPVLAQKQAL 147

RESULT 9  
 Q874E7 PRELIMINARY; PRT; 333 AA.  
 AC Q874E7;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE MAT1-2.  
 GN Name=MAT1-1-2;  
 OS Paecilomyces tenuipes.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Paecilomyces.  
 OX NCBI\_TaxID=45847;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=15321663; DOI=10.1016/j.femsle.2004.06.033;  
 RA Yokoyama E., Yamagishi K., Hara A.;  
 RT "Development of a PCR-based mating-type assay for Clavicipitaceae.";  
 RL FEMS Microbiol. Lett. 237:205-212 (2004).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22783895; PubMed=12902305;  
 RX DOI=10.1128/AEM.69.8.5019-5022.2003;  
 RA Yokoyama E., Yamagishi K., Hara A.;  
 RT "Structures of the mating-type loci of Cordyceps takamontana.";  
 RL Appl. Environ. Microbiol. 69:5019-5022 (2003).  
 DR EMBL; AB096216; BAC67540.1; -  
 SQ SEQUENCE 333 AA; 37756 MW; 5590468907FEBBF6 CRC64;

Query Match 40.6%; Score 56; DB 2; Length 333;  
 Best Local Similarity 44.8%; Pred. No. 59;  
 Matches 13; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

Qy 2 VEKIKTVKESATEEKLTPVLAQKQAL 30  
 Db 66 VNHIVDRIVERFSAEKIDPVILVKEITAL 94

## RESULT 10

Q72JL4 PRELIMINARY; PRT; 445 AA.  
 AC Q72JL4;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Biotin carboxylase (BC 6.3.4.14).  
 GN OrderedClausenases-TC0758;  
 OS Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).  
 OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;  
 OC Thermus.  
 OX NCBI\_TaxID=262724;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=15064768;  
 RA Henne A., Brueggemann H., Raasch C., Wierzer A., Hartech T.,  
 RA Liesegang H., Johann A., Lienard T., Gohl O., Martinez-Arias R.,  
 RA Jacobi C., Starkuviene V., Schlenceczek S., Dencker S., Huber R.,  
 RA Klenk H.-P., Kramer W., Merkl R., Gottschalk G., Fritz H.-J.;  
 RT "The genome sequence of the extreme thermophile Thermus  
 RT thermophilus.";  
 RL Nat. Biotechnol. 22:547-553 (2004).  
 DR EMBL; AB017303; AAS81104.1; -  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004075; F:biotin carboxylase activity; IEA.  
 DR GO; GO:0016874; F:ligase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR004549; AccC.  
 DR InterPro; IPR005482; Biotin\_carb\_C.  
 DR InterPro; IPR005481; Cphae\_L\_N.  
 DR InterPro; IPR005479; Cphae\_synth\_L\_D2.  
 DR InterPro; IPR011054; Rudmt\_hyb\_motif.  
 DR Pfam; Pf02785; Biotin\_carb\_C; 1.

```
DR Pfam; PF00289; CPSase_L_chain; 1.
DR Pfam; PF02786; CPSase_L_D2; 1.
DR TIGRFAMs; TIGR00514; accC; 1.
DR PROSITE; PS00866; CPSASE_1; UNKNOWN_1.
DR PROSITE; PS00867; CPSASE_2; UNKNOWN_1.
DR Complete proteome.
SQ SEQUENCE 445 AA; 49302 MW; B57FC8E914481A5 CRC64;

Query March 39.1%; Score 54; DB 2; Length 445;
Best Local Similarity 54.5%; Pred. No. 1.4e+02;
Matches 12; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 5 IKTVKESATEBKLPVLLAKQ 26
Db 25 IKTVVASTADEKSLPVLLADE 46

RESULT 11
ID C72Y ARATH STANDARD; PRT; 500 AA.
AC 09LIP3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Cytochrome P450 71B37 (BC 1.14.-.-).
GN Name=CYP71B37; OrderedLocustNames=At3g26330; ORFNames=F20C19.5;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eustoids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20363099; PubMed=10907853;
RA Kaneke T., Katoh T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the 4,251,695 bp regions covered by 90 Pl, TAC
RT and BAC clones."
RL DNA Res. 7:217-221(2000).
RN [2]
RP CONCEPTUAL TRANSLATION.
RP Bak S., Paquette S.;
RL Unpublished observations (APR-2001).
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to erroneous
CC gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).
CC -----
DR EMBL; AP001298; BAB02193.1; ALT_SRQ.
DR HSSP; P00179; IDT6.
DR GeneFarm; 1244; 94.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002401; EP4501.
DR Pfam; PF00067; P450.1.
DR PRINTS; PR00463; EP4501.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450.1.
DR Heme; Monooxygenase; Multigene family; Oxidoreductase; Transmembrane.
FT TRANSMM 2 22
FT METALM 440 440 Iron (heme axial ligand) (By similarity).
SQ SEQUENCE 500 AA; 56843 MW; B36088363D918FE1 CRC64;

Query Match 39.1%; Score 54; DB 1; Length 500;
Best Local Similarity 43.3%; Pred. No. 1.6e+02;
Matches 13; Conservative 9; Mismatches 6; Indels 2; Gaps 1;
```

```
Qy 1 EVEKITVKSATEBKLPVLLAKQLAAL 30
Db 150 EVKTLNMFESBAQR--TEVNLSEKASL 177

RESULT 12
ID 06BPD3 PRELIMINARY; PRT; 769 AA.
AC 06BPD3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Debaryomyces hanseni chromosome E of strain CBS767 of Debaryomyces
DE hanseni.
GN ORFNames=DEHNOE150929;
OS Debaryomyces hanseni CBS767.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX NCBI_TaxID=284592;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul I., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Bianchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boissarme A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
RA Despons L., Fabre E., Falthed C., Ferry-Dumazet H., Groppi A.,
RA Hantreaye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Niclaud J.M., Nikoleki M., Oztas S., Ozler-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swemene D., Tekaya F., Wesolowski-Jouvel M., Westhof E., Wirth B.,
RA Zenioui-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thiery A.,
RA Boulcher P., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in Yeasts."
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RG Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382137; CAG8181.1; -.
DR GO; GO:0005525; P:GTP binding; IEA.
DR GO; GO:0003746; P:translation elongation factor activity; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR GO; GO:0006414; P:translational elongation; IEA.
DR InterPro; IPR004540; EF-G.
DR InterPro; IPR000640; EFG_C.
DR InterPro; IPR009022; EFG_TII-V.
DR InterPro; IPR009022; EFG_IV.
DR InterPro; IPR005517; EFG_IV.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR000795; Protsyn_GTPbind.
DR InterPro; IPR005225; Small_GTP.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00679; EFG_C; 1.
DR Pfam; PF00764; EFG_IV; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR TIGRFAMs; TIGR00484; EF-G; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR GTP-binding; Protein biosynthesis.
SQ SEQUENCE 769 AA; 85013 MW; CA2EBCF3A8EBCC CRC64;

Query Match 39.1%; Score 54; DB 2; Length 769;
Best Local Similarity 37.0%; Pred. No. 2.4e+02;
Matches 10; Conservative 6; Mismatches 11; Indels 0; Gaps 0;
```

```

Db          311 VEOIKAIRATRIARRFTPTVLMGSLA 337
|||||:::~::~|||:~|
RESULT 13
ID GLND_AZOBR STANDARD; PRT; 933 AA.
AC OBR0D1;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE [Protein-Pil] uridylyltransferase (EC 2.7.7.59) [Pil uridylyl-
DE transferase] (Uridylyl removing enzyme) (Urase).
GN Name-glnD;
OS Azospirillum brasiliense.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Rhodospirillaceae; Azospirillum.
OX NCBI_TaxID=192;
[1]
RP SEQUENCE FROM N.A..
RC STRAIN=Sp7 / ATCC 29145;
RA Van Dommelen A., Keljers V., Somers E., Vanderleyden J.;
RT "Cloning and characterization of the Azospirillum brasiliense glnD gene
RT and analysis of a glnD mutant.";
RL Mol. Gen. Genet. 266:813-820(2002).
CC -1- FUNCTION: Modifies, by uridylylation or deuridylylation the Pil
CC (glnB) regulatory protein.
CC -1- CATALYTIC ACTIVITY: UTP + [protein-Pil] = diphosphate + uridylyl-
CC [protein-Pil].
CC -1- SIMILARITY: Belongs to the glnD family.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements
CC or send an email to licenses@ebi.ac.uk).
CC CC
DR EMBL: AF149716; AAL87737.1; -.
DR HAMAP: MF 00277; -1.
DR InterPro: IPR002912; ACT.
DR InterPro: IPR006674; HD_hydro.
DR InterPro: IPR003607; Met_phos_hydro.
DR InterPro: IPR002934; NTP_transf.
DR InterPro: IPR010043; Urase_glnD.
DR Pfam: PF01842; ACT; 2.
DR Pfam: PF01966; HD; 1.
DR Pfam: PF01909; NTP_transf_2; 1.
DR SMART: SMO0471; HDC; 1.
DR TIGRFAMs: TIGR01693; Urase_glnD; 1.
KW Nitrogen fixation; Nucleotidyltransferase; Transferase.
SQ SEQUENCE 933 AA; 104601 MW; 35E652ED0411BB9E7 CRC64;

Query Match 39.1%; Score 54; DB 1; Length 933;
Best Local Similarity 42.9%; Pred.No. 2.8e+02;
Matches 12; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

Oy 3 EKIKTYVESATEKLTPTVLAKQLAL 30
DB 36 EDLETVAEHGTGDKLRLPALIARLGAL 63

RESULT 14
ID Q6BFZ2 PRELIMINARY; PRT; 416 AA.
AC Q6BFZ2;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Hypothetical protein.
DE ORFNames=PTMB.230;
OS Paramexium tetraurelia.
```

OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculiida;  
OC Paramacium.  
OX NCBI\_TaxID=5688;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Stoek d4-2;  
RX PubMed=15926759; DOI=10.1016/j.cub.2004.07.029;  
RA Zagulski M., Nowak J.K., Le Mouel A., Nowacki M., Migdalski A.,  
RA Gromadka R., Noel B., Blanc I., Dessen P., Wincker P., Keller A.M.,  
RA Cohen U., Meyer E., Sperling L.;  
RT "High Coding Density on the largest Paramacium tetraurelia Somatic  
RT Chromosome";  
RL Curr. Biol. 14:1397-1404(2004).  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Stock d4-2;  
RX Nowak J.K., Migdalski A., Gromadka R., Zagulski M.;  
RT "Paramacium megabase sequencing project."  
RL Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; CR548612; CAHQ3428.1; -;  
DR InterPro; IPRO01251; CRAI\_TRIO\_C.  
DR PROSITE; PS50191; CRAI\_TRIO; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 416 AA; 48774 MM; 883270AAB2910D1 CRC64;

Gy 2 VERIKTVKESATEEKLTPVILAKQLAA 29

ID ID PRELIMINARY; PRF; 710 AA.

AC Q6CF32;  
DT 25-OCT-2004 (TrEMBLrel\_28, Created)  
DT 25-OCT-2004 (TrEMBLrel\_28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel\_28, Last annotation update)  
DE Similarity (Fragment).  
OE ORFName=YAL10B10692g;  
OS Yarrowia lipolytica CLIB99.  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Dipodascaceae; Yarrowia.  
OX NCBI\_TaxID=284591;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CLIB99;  
RG Genoluvures;  
RA Lafontaine I., Sherman D., Fischer G., Durrens P., Casarego A.S.,  
RA Dujon B., Dubourg J., Montigny J., Marck C., Neugeglisse C., Talla E.,  
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
RA Batnay S., Bianchin S., Beckerich J.M., Beyne E., Bleyssest C.,  
RA Boissiere A., Boyer J., Cattoletti L., Confiantolexi F., de Darovar A.,  
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,  
RA Hartave F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,  
RA Nicolaud J.M., Nikolaki M., Oztas S., Ozler-Kalogeropoulos O.,  
RA Pellenz S., Potter S., Richard G.F., Straud M.L., Suleau A.,  
RA Swemane D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,  
RA Zeniou-Meyer M., Zivanovic I., Bolocin-Fukuhara M., Thierry A.,  
RA Bouchier C., Caudeiron B., Scarpelli C., Gallardin C., Weissenbach J.,  
Wincker P., Soulier J.L.;  
RL "Genome evolution in yeasts.";  
NT Nature 430:35-44(2004).  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CLIB99;  
RA GenomeScope;  
RL Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; CR382128; CAG82975.1; -;

DR InterPro; IPR005011; SART\_1.  
 Pfam; PF03343; SART-1; 1.  
 FT NON TRR  
 SQ SEQUENCE 710 AA; 81568 MW; 89C3CA466BB2F69B CRC64;  
 Query Match 38.4%; Score 53; DB 2; Length 710;  
 Best local similarity 50.0%; Pred. No. 2.9e+02;  
 Matches 12; Conservative 3; Mismatches 9; Indels 0; Gaps 0;  
 Qy 6 KTYKESATEKLTPTVLAKQLAA 29  
 Db 426 KNLMKSQRNRKLTPEQLAEQLAA 449

Search completed: November 3, 2005, 22:03:01  
 Job time : 55.7808 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 3, 2005, 21:32:33 ; Search time 71.1872 Seconds  
(without alignments)  
282.516 Million cell updates/sec

Title: 09782816-51

Sequence: 1 GVKETPQOKTQRLHEVQEL.....ESATEEKLTPLYLAKQLAAL 52

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

Result No.	Score	Query Match	Length	ID	Description
1	251	100.0	314	8 ABM81421	Abm81421 Tumour-as
2	251	100.0	378	6 ABP98851	ABP98851 Human str
3	251	100.0	401	7 ADJ69563	ADJ69563 Human hea
4	251	100.0	406	5 ABP53018	ABP53018 Human p50
5	251	100.0	465	3 AAB58668	Aab58668 Breast an
6	248	98.8	183	5 ABP53019	ABP53019 Mouse p50
7	246	98.0	52	5 ABP53016	ABP53016 Cellular
8	231	92.0	134	4 AAG75956	Aag75956 Human col
9	113	45.0	22	5 ABP52968	ABP52968 Cellular
10	107	42.6	21	5 ABP52969	ABP52969 Cellular
11	103	41.0	20	5 ABP52970	ABP52970 Cellular
12	103	41.0	23	5 ABP52966	ABP52966 Cellular
13	98	39.0	19	5 ABP52971	ABP52971 Cellular
14	93	37.1	18	5 ABP52972	ABP52972 Cellular
15	88	35.1	17	5 ABP52973	ABP52973 Cellular
16	81	32.3	16	5 ABP52974	ABP52974 Cellular
17	76	30.3	15	5 ABP52975	ABP52975 Cellular
18	72.5	28.9	53	5 ABP53017	ABP53017 Cellular
19	72.5	28.9	380	5 ABB59088	ABBS9088 Drosophi
20	72.5	28.9	380	5 ABP53020	ABP53020 Drosophi
21	71	28.3	14	5 ABP52976	ABP52976 Cellular
22	70	27.9	1087	4 ABB65359	ABBS6359 Drosophi
23	70	27.9	1087	4 ABB65358	ABBS6358 Drosophi
24	68	27.1	2368	4 AAU34139	AAU34139 Staphyloc
25	68	27.1	2368	4 AAU36796	AAU36796 Staphyloc

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

26	67	26.7	1448	6 ADA89551	AdA89551 Staphyloc
27	67	26.7	2478	4 AAU37374	Aau37374 Staphyloc
28	67	26.7	2478	4 AAU34320	Aau34320 Staphyloc
29	67	26.7	2478	6 ABJ19002	Abj19002 Pathogen
30	67	26.7	2478	6 ABM71899	ABM71899 Staphyloc
31	67	26.7	2481	6 ABU15838	ABU15838 Protein e
32	67	26.7	2481	7 ABR62804	ABR62804 Methicill
33	66	26.3	13	5 ABP52977	ABP52977 Cellular
34	66	26.3	386	2 ABU44043	ABU44043 Protein e
35	66	26.3	868	2 AAY37731	Aay37731 Protein i
36	66	26.3	1020	4 AAM79875	Aam79875 Human pro
37	65.5	26.1	207	4 AAB29754	Aab29754 Rice pnt-
38	65.5	26.1	1837	8 ADS44304	ADs44304 Bacterial
39	64	25.5	140	4 AAO12168	AAo12168 Human pol
40	64	25.5	151	3 AAG03725	Aag03725 Human sec
41	64	25.5	188	2 AAW64545	AAw64545 Human sco
42	64	25.5	188	3 AAB10278	Aab10278 Human fet
43	64	25.5	188	4 AAU12172	AAU12172 Human PRO
44	64	25.5	188	6 ABO17616	ABO17616 Novel hum
45	64	25.5	188	6 ABU80870	ABU80870 Human PRO

## ALIGNMENTS

RESULT 1  
ABM81421  
ID ABM81421 standard; protein; 314 AA.

XX	ABM81421;	18-NOV-2004 (first entry)	Tumour-associated antigenic target (TAT) polypeptide PRO82191, SEQ:3675.
XX	AC	ABM81421;	Tumour-associated antigenic target; TAT; human; overexpression; cancer;
XX	DE	18-NOV-2004 (first entry)	lunour; diagnosis; cell proliferative disorder; breast cancer;
XX	KW	Tumour-associated antigenic target; TAT; human; overexpression; cancer;	colorectal cancer; lung cancer; ovarian cancer; liver cancer;
XX	KW	central nervous system cancer; bladder cancer; pancreatic cancer;	central nervous system cancer; melanoma; leukaemia; hybridisation probe;
XX	KW	chromosome identification; chromosome mapping; gene mapping;	gene therapy; cytostatic.
XX	KW	gene therapy; cytostatic.	
XX	OS	Homo sapiens.	
XX	OS		
XX	PN	WO2004030615-A2.	
XX	XX	15-APR-2004.	
XX	PD	29-SEP-2003; 2003WO-US028547.	
XX	PE	02-OCT-2002; 2002US-0414971P.	
XX	PR	(GERTH ) GENENTECH INC.	
XX	PA	Wu TD, Zhang Z, Zhou Y;	
XX	PI	WPI: 2004-347921/32.	
XX	PI	N-PSDB; ACN39497.	
XX	DR	New tumor-associated antigenic target polypeptides and nucleic acids,	
XX	DR	useful in preparing a medicament for treating or detecting a	
XX	XX	proliferative disorder, e.g. breast, lung, colorectal, ovarian or	
XX	XX	prostate cancer or tumor.	
XX	PS	Claim 12; SEQ ID NO 3675; 7273bp; English.	
XX	XX	The invention relates to human tumour-associated antigenic target (TAT)	
XX	XX	polypeptides, and their related nucleic acids. The TAT polypeptides are	
XX	CC	overexpressed in cancer tissues compared to normal tissues, and may thus	
XX	CC	serve as effective targets for the diagnosis and treatment of cancer in	
XX	CC	mammals. The invention also relates to nucleic acid and polypeptide	
XX	CC	sequences at least 80% identical to the TAT nucleic acids and	

Cc	polypeptides; expression vectors and host cells comprising a TAT nucleic acid;
Cc	an antibody specific for a TAT polypeptide; a peptide or organic molecule which binds to a TAT polypeptide; fusion proteins comprising a TAT polypeptide; and methods and compositions for the treatment or diagnosis of cancer in mammals.
Cc	TAT polypeptides, nucleic acids, antibodies, antagonists, binding molecules and compositions are useful for diagnosing or treating a cell proliferative disorder associated with increased TAT expression, particularly cancers such as breast cancer, colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder cancer, pancreatic cancer, cervical cancer, cancers of the central nervous system, melanoma and leukemia. TAT nucleic acids may further be used as hybridisation probes, in chromosome and gene mapping, in chromosome identification and in gene therapy. The present sequence represents a TAT polypeptide of the invention
Cc	
SQ	Sequence 314 AA:
Oy	Query Match            100.0%   Score 251; DB 8; Length 314; Best Local Similarity   100.0%; Pred No.6e-21; Matches         52; Conservative      0; Mismatches          0; Indels        0; Gaps            0  
Dd	7 GVKEKPPQQKYQRLLHEVDELTEVEKIKTTVKSATESEEKITPVLAKQLAAL 52           GVKEKPPQQKYQRLLHEVDELTEVEKIKTTVKSATESEEKITPVLAKQLAAL 58  
ID	ABP98851 standard; protein; 378 AA.  
AC	ABP98851;  
DT	15-JUL-2003 (first entry)  
DE	Human structural and cytoskeletal associated protein #42.  
KX	Cytostatic; antiarteriosclerotic; anticonvulsant; nootropic; antianginal; neuroprotective; cerebroprotective; hypotensive; cardiact; osteopathic; antiinflammatory; antiarthritic; virucide; gene therapy; human; stroke; structural and cytoskeleton-associated protein; SCAP; cancer; angina;therosclerosis; epilepsy; Huntingdon's disease; hypertension; heart failure; osteoporosis; osteoarthritis.  
KM	Homo sapiens.  
OS	WO2003031940-A2.  
PN	17-Apr-2003.  
PD	10-OCT-2002; 2002WO-US032851.  
PF	12-OCT-2001; 2001US-0328931P. 19-OCT-2001; 2001US-0350681P. PR 02-Nov-2001; 2001US-0343896P. PR 09-Nov-2001; 2001US-0346308P. PR 16-Nov-2001; 2001US-0332385P. PR 07-Dec-2001; 2001US-0340776P. PR 11-Jan-2002; 2002US-0347703P.  
XX	(INCY-) INCYTE GENOMICS INC.  
PA	Bachia SD, Bhatia UJ, Borowsky ML, Butrill JD, Chang H, PI Bechla NK, Elliott VS, Emerging BM, Forsythe IY, Gorvad AE; PI Griffin JA, Hafalia AJA, Ho A, Isen CH, Kable AF, Khare R, Lal PG; PI Lee S, Lee BA, Lee SY, Leht-Nason PM, Li JX, Lindquist EA, Luo W; PI Marquis JP, Ramkumar Y, Richardson TW, Sprague MW, Swarnakar A; PI Tang YT, Warren BA, Yang J, Yue H, Zebardjadian Y, Zheng W;  
DR	MPI: 2003-403125/38.  
N-	N-PSTDB; ACC44338.  
PT	New human structural and cytoskeleton-associated proteins (SCAP) useful for diagnosing, treating and preventing diseases or conditions associated

PT	with the aberrant SCAP expression e.g. cancer, osteoporosis, or epilepsy.
XX	
XX	Claim 1; Page 298; 361pp; English.
XX	
CC	This sequence represents a novel isolated human structural and
CC	cytoskeleton-associated protein (SCAP) polypeptide. The polypeptides and
CC	polynucleotides encoding them are useful in diagnosing, treating and
CC	preventing diseases or conditions associated with the decreased
CC	expression or over expression of SCAP, such as cell proliferative (e.g.
CC	cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's
CC	disease, stroke), heart (e.g. hypertension, heart failure, angina) and
CC	skeletal muscle disorders (e.g. osteoporosis, osteoarthritis) or viral
CC	infections. These are also useful in assessing the effects of exogenous
CC	compounds on the expression of nucleic acid and amino acid sequences of
CC	SCAP. The SCAP or its fragments are useful in screening compounds for
CC	effectiveness as agonist or antagonist of the polypeptides, or in
CC	altering the expression of the target polynucleotide and compounds that
CC	specifically bind to or modulate the activity of the polypeptide. The
CC	microarray is useful in monitoring or measuring protein-protein
CC	interactions, drug-target interactions, and gene expression profiles
XX	
SQ	Sequence 378 AA;
Query Match	100.0%; Score 251; DB 6; Length 378;
Best Local Similarity	100.0%; Pred. NO.7.4e-21;
Matches	52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 GVKETPOOKYORLLHEVOELTFTEYEKIKTYVKSATSEKLTPLVLLAKQLAL 52
Db	71 GVKETPOOKYORLLHEVOELTFTEYEKIKTYVKSATSEKLTPLVLLAKQLAL 122
RESULT 3	
ID	ADJ69563
XX	ADJ69563 standard; protein; 401 AA.
XX	
ADJ69563;	
DT	
XX	06-MAY-2004 (first entry)
XX	
DE	Human heat mitochondrial protein as a therapeutic target SeqIDJ369.
XX	
KW	mitochondrial; human; screening assay; diabetes mellitus;
KW	Huntington's disease; osteoarthritis;
KW	Leber's hereditary Optic neuropathy; LHON;
KW	mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW	myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW	neuroprotective; nocotropic; antidiabetic; anticonvulsant; antiarthritic;
KW	osteopathic; ophthalmological; cyostatic.
XX	
OS	Homo sapiens.
XX	
PN	WO2003087768-A2.
PD	
XX	23-OCT-2003.
XX	
PF	04-APR-2003; 2003WO-US010870.
XX	
PR	12-APR-2002; 2002US-0372843P.
PR	17-UN-2002; 2002US-0389987P.
PR	20-SEP-2002; 2002US-0412418P.
XX	
XX	
PA	(MITO-) MITOKOR.
PA	(BUCK-) BUCK INST AGE RES.
XX	
PI	Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GW,
XX	Warnock DB;
DR	WPI; 2003-045369/78.
XX	
PT	Identifying a mitochondrial target for drug screening assays and for
PT	treating diseases associated with altered mitochondrial function,
PT	comprises detecting a modified polypeptide in a sample and correlating

PT with the disease.  
XX  
PS Claim 1; SEQ ID NO 1369; 180pp; English.  
CC This invention relates to novel mitochondrial targets that can be used  
CC for therapeutic intervention in treating a disease associated with  
CC altered mitochondrial function. Specifically, it refers to a method for  
CC identifying proteins of the human heart mitochondrial proteome that are  
CC useful for drug screening assays, as well as therapeutic targets. The  
CC present invention describes a method for identifying such proteins that  
CC can be used in the treatment of various diseases associated with altered  
CC mitochondrial function including diabetes mellitus, Huntington's disease,  
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial  
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy  
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these  
CC compositions have neuroprotective, neurotropic, antidiabetic,  
CC anticonvulsant, antiarrhythmic, osteopathic, ophthalmological and  
CC cyostatic activities. This polypeptide sequence is a human heart  
CC mitochondrial protein of the invention.  
SQ Sequence 401 AA;  
Query Match 100.0%; Score 251; DB 7; Length 401;  
Best Local Similarity 100.0%; Pred. No. 7,9e-21;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 GVKETPOOKYORLHVEVQLTTEVEKIKTVKESATBEKLPVLLAKQLAAL 52  
DB 94 GVKETPOOKYORLHVEVQLTTEVEKIKTVKESATBEKLPVLLAKQLAAL 145  
RESULT 4  
ABP53018  
ID ABP53018 standard; protein; 406 AA.  
XX  
AC ABP53018;  
XX  
DT 05-NOV-2002 (first entry)  
XX  
DE Human p50 amino acid sequence SEQ ID NO:53.  
XX  
KW Cellular proliferation inhibition; cytostatic; antiinflammatory; cancer;  
KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;  
KW glandular disorder; macropagal disorder; epithelial disorder;  
KW stromal disorder; blastococelic disorder; angiogenic disorder;  
KW immunologic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200264779-A2.  
XX  
PD 22-AUG-2002.  
XX  
PE 21-JAN-2002; 2002WO-US001708.  
XX  
PR 14-FEB-2001; 2001US-00782816.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX  
PI Sharp DJ, Rogers GC, Scholey JM;  
XX  
DR WPI; 2002-657599/70.  
XX  
PT New peptide inhibitors of p50/dynaminin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.  
XX  
PS Disclosure; Fig 1; 55pp; English.  
XX  
CC The present invention describes an isolated peptide (I) comprising or

CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
CC terminal extensions. (I) have cyostatic and antiinflammatory activities  
CC and can be used as p50/dynaminin inhibitors and in gene therapy. The  
CC peptides, nucleic acid molecules and methods from the present invention  
CC are useful for treating cancer by inhibiting cellular proliferation, such  
CC as benign or malignant tumours (renal, liver, kidney, bladder, breast,  
CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
CC neck tumours); leukemias and lymphoid malignancies, other disorders such  
CC as neuronal, glial, astrocytal, hypothalamic and other glandular,  
CC macropagal, epithelial, stromal and blastococelic disorders; and  
CC inflammatory, angiogenic and immunologic disorders. The present sequence  
CC represents human p50 which is given in the exemplification of the present  
CC invention  
SQ Sequence 406 AA;  
Query Match 100.0%; Score 251; DB 5; Length 406;  
Best Local Similarity 100.0%; Pred. No. 8,1e-21;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 GVKETPOOKYORLHVEVQLTTEVEKIKTVKESATBEKLPVLLAKQLAAL 52  
DB 99 GVKETPOOKYORLHVEVQLTTEVEKIKTVKESATBEKLPVLLAKQLAAL 150  
RESULT 5  
AAB58968  
ID AAB58968 standard; protein; 465 AA.  
XX  
AC AAB58968;  
XX  
DT 27-MAR-2001 (first entry)  
XX  
DE Breast and ovarian cancer associated antigen protein sequence SRQ ID 676.  
XX  
KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;  
KW neurotropic; neuroprotective; antiviral; antiatherogenic; hepatotropic;  
KW antidiabetic; antiinflammatory; antitumor; vulnerary; anticonvulsant;  
KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;  
KW Addison's disease; allergy; autoimmune haemolytic anemia;  
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;  
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;  
KW cardiovascular disorder; wound healing; neurological disease.  
XX  
OS Homo sapiens.  
XX  
PN WO20005173-A1.  
XX  
PD 21-SEP-2000.  
XX  
PE 08-MAR-2000; 2000WO-US005881.  
XX  
PR 12-MAR-1999; 99US-0124270P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Ruben SM;  
XX  
DR WPI; 2000-611515/58.  
XX  
PT N-PSDB; AAF21871.  
XX  
PT New human breast and ovarian cancer associated gene sequences and the  
PT polypeptides encoded by these genes, useful in the prevention, treatment  
PT and diagnosis of cancer, immune disorders, cardiovascular disorders and  
PT neurological diseases.  
XX  
PS Claim 11; Page 1126-1128, 1299pp; English.  
XX  
CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human  
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are  
CC associated with breast and ovarian cancer. Included in the invention are



CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the  
CC isolation and characterisation of the DNA and protein sequences of the  
CC invention. The breast and ovarian cancer associated DNA, protein, agonist  
CC or antagonist sequences exhibit cytostatic, immunosuppressive, nocotropic,  
CC neuroprotective; antiviral; antiallergic; hepatocytic; antidiabetic;  
CC antiinflammatory; antitumor; anticonvulsant; antibacterial;  
CC antifungal; antiparasitic and cardiant activity. The polynucleotide and  
CC protein sequences are used in the diagnosis of cancer, particularly  
CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists  
CC and agonists may also be used in the diagnosis, prevention and treatment  
CC of immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
CC cardiovascular disorders such as myocardial ischaemias; wound healing;  
CC neurological diseases such as cerebral anoxia and epilepsy; and  
CC infectious diseases  
CC  
CC Sequence 465 AA;  
SQ

Query Match 100.0%; Score 251; DB 3; Length 465;  
Best Local Similarity 100.0%; Pred. No. 9,4e-21;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVKETPOOKYORLLHEVQELTTEVEKIKTYVESATEBEKLTPLVLAQQLAAL 52  
|||  
158 GVKETPOOKYORLLHEVQELTTEVEKIKTYVESATEBEKLTPLVLAQQLAAL 209

Db 158 GVKETPOOKYORLLHEVQELTTEVEKIKTYVESATEBEKLTPLVLAQQLAAL 209

RESULT 6  
AAB53019  
ID AAB53019 standard; protein; 183 AA.  
XX  
AC AAB53019;  
XX  
DT 05-NOV-2002 (first entry)  
XX  
XX Mouse p50 amino acid sequence SEQ ID NO:54.  
XX  
XX Cellular proliferation inhibition; cytostatic; antiinflammatory; cancer;  
KM p50 inhibitor; dynaminin inhibitor; gene therapy; tumor; carcinoma;  
KM sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
KM glial disorder; astrocytal disorder; hypochalamic disorder; inflammatory;  
KM glandular disorder; macrophagal disorder; epithelial disorder;  
KM stromal disorder; blastococelic disorder; angiogenic disorder;  
KM immunologic disorder.  
XX  
XX Mus musculus.  
OS  
XX WO200264779-A2.  
PN  
XX 22-AUG-2002.  
PD  
XX 21-JAN-2002; 2002WO-US001708.  
PF  
XX 14-FEB-2001; 2001US-00782816.  
PR  
XX (REGC ) UNIV CALIFORNIA.  
PA  
XX Sharp DJ, Rogers GC, Scholey JM;  
PI  
XX WPI; 2002-657599/70.  
DR  
XX  
XX New peptide inhibitors of p50/dynaminin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.  
XX  
XX Discloure; Fig 2; 55pp; English.  
XX  
CC The present invention describes an isolated peptide (I) comprising or  
CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
CC sequences given in AAB52966 and AAB52967 and can have C-terminal and N-  
CC terminal extensions. (I) have cytostatic and antiinflammatory activities

CC and can be used as p50/dynaminin inhibitors and in gene therapy. The  
CC peptides, nucleic acid molecules and methods from the present invention  
CC are useful for treating cancer by inhibiting cellular proliferation, such  
CC as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
CC neck tumors); leukaemias and lymphoid malignancies, other disorders such  
CC as neuronal, glial, astrocytal, hypochalamic and other glandular,  
CC macrophagal, epithelial, stromal and blastococelic disorders; and  
CC inflammatory, angiogenic and immunologic disorders. The present sequence  
CC represents mouse p50 which is given in the exemplification of the present  
CC invention  
CC  
CC Sequence 183 AA;  
SQ

Query Match 98.8%; Score 248; DB 5; Length 183;  
Best Local Similarity 98.1%; Pred. No. 7,1e-21;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVKETPOOKYORLLHEVQELTTEVEKIKTYVESATEBEKLTPLVLAQQLAAL 52  
|||  
94 GVKETPOOKYORLLHEVQELTTEVEKIKTYVESATEBEKLTPLVLAQQLAAL 145

Db 94 GVKETPOOKYORLLHEVQELTTEVEKIKTYVESATEBEKLTPLVLAQQLAAL 145

RESULT 7  
AAB53016  
ID AAB53016 standard; peptide; 52 AA.  
XX  
AC AAB53016;  
XX  
DT 05-NOV-2002 (first entry)  
XX  
XX Cellular proliferation peptide inhibitor SEQ ID NO:51.  
XX  
XX  
KM Cellular proliferation inhibition; cytostatic; antiinflammatory; cancer;  
KM p50 inhibitor; dynaminin inhibitor; gene therapy; tumor; carcinoma;  
KM sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
KM glial disorder; astrocytal disorder; hypochalamic disorder; inflammatory;  
KM glandular disorder; macrophagal disorder; epithelial disorder;  
KM stromal disorder; blastococelic disorder; angiogenic disorder;  
KM immunologic disorder.  
XX  
XX Homo sapiens.  
OS  
XX Mus musculus.  
OS  
XX Synthetic.  
OS  
XX Key Location/Qualifiers  
FH Misc-difference 44 /label= 'Leu, Val  
FT /note= "Leu in humans and Val in Mus musculus"  
FT  
XX  
XX WO200264779-A2.  
PN  
XX 22-AUG-2002.  
PD  
XX 21-JAN-2002; 2002WO-US001708.  
PF  
XX 14-FEB-2001; 2001US-00782816.  
PR  
XX (REGC ) UNIV CALIFORNIA.  
PA  
XX Sharp DJ, Rogers GC, Scholey JM;  
PI  
XX WPI; 2002-657599/70.  
DR  
XX  
XX New peptide inhibitors of p50/dynaminin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.  
XX  
XX Claim 2; Page 31; 55pp; English.  
XX  
CC The present invention describes an isolated peptide (I) comprising or

CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
CC terminal extensions. (1) have cytostatic and antiinflammatory activities  
CC and can be used as p50/dynaminin inhibitors and in gene therapy. The  
CC peptides, nucleic acid molecules and methods from the present invention  
CC are useful for treating cancer by inhibiting cellular proliferation, such  
CC as benign or malignant tumours (renal, liver, kidney, bladder, breast,  
CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
CC neck tumours); leukemias and lymphoid malignancies, other disorders such  
CC as neuronal, glial, astrocytal, hypochalamic and other glandular,  
CC macrophagal, epithelial, stromal and blastocoele disorders; and  
CC inflammatory, angiogenic and immunologic disorders. The present sequence  
CC represents a specifically claimed peptide inhibitor of cellular  
CC proliferation from the present invention

XX  
XX  
SQ Sequence 52 AA;

Query Match 98.0%; Score 246; DB 5; Length 52;  
Best Local Similarity 98.1%; Pred. No. 2.8e-21;  
Matches 51; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVKETPOOKYQRLHEVOELTTEVEKIKTYVESATEERKLPVLLAKQALAL 52  
Db 1 GVKETPOOKYQRLHEVOELTTEVEKIKTYVESATEERKLPVLLAKQALAL 52

RESULT 8  
AAG75956  
ID AAG75956 standard; protein; 134 AA.

XX  
XX AAG75956;

XX  
XX 03-SEP-2001 (first entry)

XX  
XX Human colon cancer antigen protein SEQ ID NO:6720.

DE  
XX Human colon cancer antigen protein SEQ ID NO:6720.

KM Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KM colorectal carcinoma; chromosome 12.

XX  
XX Homo sapiens.

XX  
XX WO200122920-A2.

XX  
XX 05-APR-2001.

XX  
XX 28-SEP-2000; 2000WO-US026524.

XX  
XX 29-SEP-1999; 99US-0157137P.  
XX 03-NOV-1999; 99US-0163280P.

PR  
XX (HUMA-) HUMAN GENOME SCI INC.

XX  
XX Ruben SM, Barash SC, Birse CE, Rosen CA;  
XX WPI; 2001-235357/24.  
XX N-PSDB; AAH35361.

DR  
XX

XX  
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
XX useful for preventing, diagnosing and/or treating colorectal cancers.

PT  
XX

PS Claim 11; Page 8181; 9803pp; English.

XX  
XX AAH32943 to AAH37195 and AAG75956 represent human colon  
XX cancer-associated nucleic acid molecules (N) and proteins (P), where the  
XX proteins are collectively known as colon cancer antigens. The colon  
XX cancer antigens have cytostatic activity and can be used in gene therapy  
XX and vaccine production. N and P may be used in the prevention, diagnosis  
XX and treatment of diseases associated with inappropriate P expression. For  
XX example, N and P may be used to treat disorders associated with decreased  
XX expression by rectifying mutations or deletions in a patient's genome  
XX that affect the activity of P by expressing inactive proteins or to  
XX supplement the patients own production of P. Additionally, N may be used

CC to produce the colon cancer-associated Pe, by inserting the nucleic acids  
CC into a host cell and culturing the cell to express the proteins. N and P  
CC can be used in the prevention, diagnosis and treatment of colorectal  
CC carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent  
CC sequences used in the exemplification of the present invention. N.B.  
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at  
CC time of publication, meaning no sequences are present for SEQ ID NO:11027  
CC to 1052, 7921 and 7922

XX  
XX  
SQ Sequence 134 AA;

Query Match 92.0%; Score 231; DB 4; Length 134;  
Best Local Similarity 100.0%; Pred. NO. 4.8e-19;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TPQOKYQRLHEVOELTTEVEKIKTYVESATEERKLPVLLAKQALAL 52  
Db 1 TPQOKYQRLHEVOELTTEVEKIKTYVESATEERKLPVLLAKQALAL 48

RESULT 9  
ABP52968  
ID ABP52968 standard; peptide; 22 AA.

XX  
XX ABP52968;

XX  
XX 05-NOV-2002 (first entry)

XX  
XX Cellular proliferation inhibitor related peptide SEQ ID NO:3.

DE  
XX

XX  
XX Cellular proliferation inhibition; cytostatic; antiinflammatory; cancer;  
XX p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
XX sarcoma; glioblastoma; leukaemia; lymphoid malignancy; neuronal disorder;  
XX glial disorder; astrocytal disorder; hypochalamic disorder; inflammatory;  
XX glandular disorder; macrophagal disorder; epithelial disorder;  
XX stromal disorder; blastocoele disorder; angiogenic disorder;  
XX immunologic disorder.

XX  
XX Homo sapiens.

XX  
XX Mus musculus.

OS  
XX Synthetic.

XX  
XX WO200264779-A2.

XX  
XX 22-AUG-2002.

XX  
XX 21-JAN-2002; 2002WO-US001708.

XX  
XX 14-FEB-2001; 2001US-00782816.

PR  
XX (REGC ) UNIV CALIFORNIA.

XX  
XX Sharp DJ, Rogers GC, Scholey JW;  
XX WPI; 2002-657599/70.

DR  
XX

XX  
XX New peptide inhibitors of p50/dynaminin useful for treating cancer by  
XX inhibiting cellular proliferation, e.g. benign or malignant tumors,  
XX leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
XX immunologic disorders.

PT  
XX

XX  
XX Claim 1; Page 29; 55pp; English.

XX  
XX The present invention describes an isolated peptide (I) comprising or  
XX having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
XX sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
XX terminal extensions. (1) have cytostatic and antiinflammatory activities  
XX and can be used as p50/dynaminin inhibitors and in gene therapy. The  
XX peptides, nucleic acid molecules and methods from the present invention  
XX are useful for treating cancer by inhibiting cellular proliferation, such  
XX as benign or malignant tumours (renal, liver, kidney, bladder, breast,  
XX gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
XX thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and

CC neck tumours); leukemias and lymphoid malignancies, other disorders such  
CC as neuronal, glial, astrocytic, hypothalamic and other glandular,  
CC macrophagal, epithelial, stromal and blastocoeleic disorders; and  
CC inflammatory, angiogenic and immunologic disorders. The present sequence  
CC represents a peptide that can be N-terminally added to (P1)

XX Sequence 22 AA;

Query Match 45.0%; Score 113; DB 5; Length 22;  
Best Local Similarity 100.0%; Pred. No. 3.4e-06;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKETPOOKYQRLHVEQELTT 22  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 GKETPOOKYQRLHVEQELTT 22

RESULT 10  
ABP52969  
ID ABP52969 standard; peptide; 21 AA.  
XX  
AC ABP52969;

DT 05-NOV-2002 (first entry)

XX Cellular proliferation inhibitor related peptide SEQ ID NO:4.

XX Cellular proliferation inhibition; cytostatic; antiinflammatory; cancer;  
KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytic disorder; hypothalamic disorder; inflammatory;  
KW glandular disorder; macrophagal disorder; epithelial disorder;  
KW stromal disorder; blastocoeleic disorder; angiogenic disorder;  
KW immunologic disorder.

XX Homo sapiens.  
OS Mus musculus.  
OS Synthetic.

XX WO200264779-A2.

XX 22-AUG-2002.

XX 21-JAN-2002; 2002WO-US001708.

XX 14-FEB-2001; 2001US-00782816.

XX (REGC ) UNIV CALIFORNIA.

PI Sharp DJ, Rogers GC, Scholey JM;

DR WPI; 2002-657599/70.

XX New peptide inhibitors of p50/dynaminin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.

XX Claim 1; Page 29; 55pp; English.

XX The present invention describes an isolated peptide (I) comprising or  
CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
CC terminal extensions. (I) have cytostatic and antiinflammatory activities  
CC and can be used as p50/dynaminin inhibitors and in gene therapy. The  
CC peptides, nucleic acid molecules and methods from the present invention  
CC are useful for treating cancer by inhibiting cellular proliferation, such  
CC as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
CC neck tumours); leukemias and lymphoid malignancies, other disorders such  
CC as neuronal, glial, astrocytic, hypothalamic and other glandular,  
CC macrophagal, epithelial, stromal and blastocoeleic disorders; and

CC inflammatory, angiogenic and immunologic disorders. The present sequence  
CC represents a peptide that can be N-terminally added to (P1)

XX Sequence 21 AA;

Query Match 42.6%; Score 107; DB 5; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VKETPOOKYQRLHVEQELTT 22  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 VKETPOOKYQRLHVEQELTT 21

RESULT 11  
ABP52970  
ID ABP52970 standard; peptide; 20 AA.  
XX  
AC ABP52970;

DT 05-NOV-2002 (first entry)

XX Cellular proliferation inhibitor related peptide SEQ ID NO:5.

XX Cellular proliferation inhibition; cytostatic; antiinflammatory; cancer;  
KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytic disorder; hypothalamic disorder; inflammatory;  
KW glandular disorder; macrophagal disorder; epithelial disorder;  
KW stromal disorder; blastocoeleic disorder; angiogenic disorder;  
KW immunologic disorder.

XX Homo sapiens.  
OS Mus musculus.  
OS Synthetic.

XX WO200264779-A2.

XX 22-AUG-2002.

XX 21-JAN-2002; 2002WO-US001708.

XX 14-FEB-2001; 2001US-00782816.

XX (REGC ) UNIV CALIFORNIA.

PI Sharp DJ, Rogers GC, Scholey JM;

DR WPI; 2002-657599/70.

XX New peptide inhibitors of p50/dynaminin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.

XX Claim 1; Page 29; 55pp; English.

XX The present invention describes an isolated peptide (I) comprising or  
CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
CC terminal extensions. (I) have cytostatic and antiinflammatory activities  
CC and can be used as p50/dynaminin inhibitors and in gene therapy. The  
CC peptides, nucleic acid molecules and methods from the present invention  
CC are useful for treating cancer by inhibiting cellular proliferation, such  
CC as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
CC neck tumours); leukemias and lymphoid malignancies, other disorders such  
CC as neuronal, glial, astrocytic, hypothalamic and other glandular,  
CC macrophagal, epithelial, stromal and blastocoeleic disorders; and  
CC inflammatory, angiogenic and immunologic disorders. The present sequence  
XX represents a peptide that can be N-terminally added to (P1)

SQ Sequence 20 AA;  
Query Match 41.0%; Score 103; DB 5; Length 20;  
Best Local Similarity 100.0%; Pred. No. 4.5e-05;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 3 KETPOQKRYRLHVEQVLLT 22  
1 KETPOQKRYRLHVEQVLLT 20  
DB  
RESULT 12  
ABP52966  
ID ABP52966 standard; peptide; 23 AA.  
AC ABP52966;  
DT 05-NOV-2002 (first entry)  
XX Cellular proliferation inhibitor related peptide SEQ ID NO:1.  
DE  
XX Cellular proliferation inhibitor; cytostatic; antiinflammatory; cancer;  
KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukaemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;  
KW glandular disorder; macrophagal disorder; epithelial disorder;  
KW stromal disorder; blastocoelec disorder; angiogenic disorder;  
KW immunologic disorder.  
XX  
XX Homo sapiens.  
OS Mus musculus.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 22  
FT /label= 'Leu, Val  
FT /note= "Leu in humans and Val in Mus musculus"  
XX  
XX WO200264779-A2.  
XX  
XX 22-AUG-2002.  
XX  
XX 21-JAN-2002; 2002WO-US001708.  
XX  
XX 14-FEB-2001; 2001US-00782816.  
XX  
XX (REGC ) UNIV CALIFORNIA.  
XX  
PI Sharp DJ, Rogers GC, Scholey JM;  
XX  
DR WPI; 2002-657599/70.  
XX  
XX New peptide inhibitors of p50/dynaminin useful for treating cancer: by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.  
XX  
XX Claim 1; Page 29; 55pp; English.  
XX  
XX The present invention describes an isolated peptide (I) comprising or  
XX having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
XX sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
XX terminal extensions. (I) have cytostatic and antiinflammatory activities  
XX and can be used as p50/dynaminin inhibitors and in gene therapy. The  
XX peptides, nucleic acid molecules and methods from the present invention  
XX are useful for treating cancer by inhibiting cellular proliferation, such  
XX as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
XX gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
XX thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
XX neck tumors); leukaemias and lymphoid malignancies, other disorders such  
XX as neuronal, glial, astrocytal, hypothalamic and other glandular,  
XX macrophagal, epithelial, stromal and blastocoelec disorders; and  
XX inflammatory, angiogenic and immunologic disorders

XX  
SQ Sequence 23 AA;  
Query Match 41.0%; Score 103; DB 5; Length 23;  
Best Local Similarity 95.7%; Pred. No. 5.3e-05;  
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 23 EWEKIKTTVKESATEBKLTPLYL 45  
1 EWEKIKTTVKESATEBKLTPLYL 23  
DB  
RESULT 13  
ABP52971  
ID ABP52971 standard; peptide; 19 AA.  
AC ABP52971;  
DT 05-NOV-2002 (first entry)  
XX Cellular proliferation inhibitor related peptide SEQ ID NO:6.  
DE  
XX Cellular proliferation inhibitor; cytostatic; antiinflammatory; cancer;  
KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukaemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;  
KW glandular disorder; macrophagal disorder; epithelial disorder;  
KW stromal disorder; blastocoelec disorder; angiogenic disorder;  
KW immunologic disorder.  
XX  
XX Homo sapiens.  
OS Mus musculus.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 22  
FT /label= 'Leu, Val  
FT /note= "Leu in humans and Val in Mus musculus"  
XX  
XX WO200264779-A2.  
XX  
XX 22-AUG-2002.  
XX  
XX 21-JAN-2002; 2002WO-US001708.  
XX  
XX 14-FEB-2001; 2001US-00782816.  
XX  
XX (REGC ) UNIV CALIFORNIA.  
XX  
PI Sharp DJ, Rogers GC, Scholey JM;  
XX  
DR WPI; 2002-657599/70.  
XX  
XX New peptide inhibitors of p50/dynaminin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.  
XX  
XX Claim 1; Page 29; 55pp; English.  
XX  
XX The present invention describes an isolated peptide (I) comprising or  
XX having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
XX sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
XX terminal extensions. (I) have cytostatic and antiinflammatory activities  
XX and can be used as p50/dynaminin inhibitors and in gene therapy. The  
XX peptides, nucleic acid molecules and methods from the present invention  
XX are useful for treating cancer by inhibiting cellular proliferation, such  
XX as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
XX gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
XX thyroid, hepatic carcinomas, sarcomas, glioblastomas, other disorders such  
XX as neuronal, glial, astrocytal, hypothalamic and other glandular,  
XX macrophagal, epithelial, stromal and blastocoelec disorders; and  
XX inflammatory, angiogenic and immunologic disorders. The present sequence  
XX represents a peptide that can be N-terminally added to (P1)  
SQ Sequence 19 AA;

Query Match 39.0%; Score 98; DB 5; Length 19;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4 ETPQOKYORLHVEQELTT 22  
1 ETPQOKYORLHVEQELTT 19

## RESULT 14

ABP52972 ID ABP52972 standard; peptide; 18 AA.

AC ABP52972;

DT 05-NOV-2002 (first entry)

DE Cellular proliferation inhibitor related peptide SEQ ID NO:7.

KW Cellular proliferation inhibition; cytostatic; antiinflammatory; cancer;  
KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;  
KW glandular disorder; macrophagal disorder; epithelial disorder;  
KW stromal disorder; blastococelic disorder; angiogenic disorder;  
KW immunologic disorder.

OS Homo sapiens.  
OS Mus musculus.  
OS Synthetic.

PN MO200264779-A2.

PD 22-AUG-2002.

PE 21-JAN-2002; 2002MO-US001708.

PR 14-FEB-2001; 2001US-00782816.

PA (REGC ) UNIV CALIFORNIA.

PI Sharp DJ, Rogers GC, Scholey JM;

DR WPI; 2002-657599/70.

PT New peptide inhibitors of p50/dynaminin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.

PS Claim 1; Page 29; 55pp; English.

XX The present invention describes an isolated peptide (I) comprising or  
XX having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
XX sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
XX terminal extensions. (I) have cytostatic and antiinflammatory activities  
XX and can be used as p50/dynaminin inhibitors and in gene therapy. The  
XX peptides, nucleic acid molecules and methods from the present invention  
XX are useful for treating cancer by inhibiting cellular proliferation, such  
XX as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
XX gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
XX thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
XX neck tumors); leukemias and lymphoid malignancies, other disorders such  
XX as neuronal, glial, astrocytal, hypothalamic and other glandular,  
XX macrophagal, epithelial, stromal and blastococelic disorders; and  
XX inflammatory, angiogenic and immunologic disorders; and  
XX represents a peptide that can be N-terminally added to (P1)

SQ Sequence 18 AA;

Query Match 37.1%; Score 93; DB 5; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.00058;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5 TPQOKYORLHVEQELTT 22  
Db 1 TPQOKYORLHVEQELTT 18

## RESULT 15

ABP52973 ID ABP52973 standard; peptide; 17 AA.

AC ABP52973;

DT 05-NOV-2002 (first entry)

DE Cellular proliferation inhibitor related peptide SEQ ID NO:8.

KW Cellular proliferation inhibition; cytostatic; antiinflammatory; cancer;  
KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;  
KW glandular disorder; macrophagal disorder; epithelial disorder;  
KW stromal disorder; blastococelic disorder; angiogenic disorder;  
KW immunologic disorder.

OS Homo sapiens.  
OS Mus musculus.  
OS Synthetic.

PN MO200264779-A2.

PD 22-AUG-2002.

PE 21-JAN-2002; 2002MO-US001708.

PR 14-FEB-2001; 2001US-00782816.

PA (REGC ) UNIV CALIFORNIA.

PI Sharp DJ, Rogers GC, Scholey JM;

DR WPI; 2002-657599/70.

PT New peptide inhibitors of p50/dynaminin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.

PS Claim 1; Page 29; 55pp; English.

XX The present invention describes an isolated peptide (I) comprising or  
XX having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
XX sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
XX terminal extensions. (I) have cytostatic and antiinflammatory activities  
XX and can be used as p50/dynaminin inhibitors and in gene therapy. The  
XX peptides, nucleic acid molecules and methods from the present invention  
XX are useful for treating cancer by inhibiting cellular proliferation, such  
XX as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
XX gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
XX thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
XX neck tumors); leukemias and lymphoid malignancies, other disorders such  
XX as neuronal, glial, astrocytal, hypothalamic and other glandular,  
XX macrophagal, epithelial, stromal and blastococelic disorders; and  
XX inflammatory, angiogenic and immunologic disorders. The present sequence  
XX represents a peptide that can be N-terminally added to (P1)

SQ Sequence 17 AA;

Query Match 35.1%; Score 88; DB 5; Length 17;  
Best Local Similarity 100.0%; Pred. No. 0.0021;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 PQOKYORLHVEQELTT 22  
|||||

Db 1 POKKORLHEVELETT 17

Search completed: November 3, 2005, 21:57:30  
Job time : 72.1872 secs

01:21:40:00 00:00:00

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: November 3, 2005, 21:32:33 ; Search time 41.0695 Seconds  
(without alignments)  
282.516 Million cell updates/sec

Title: 09782816-1-22

Perfect score: 138  
Sequence: 1 EVEKIKTTWKESATBEKLTPTVILAKQLAAL 30  $\chi_1 = 0$

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Database : Listing first 45 summaries

A\_Geneseq\_16Dec04:\*  
1: Geneseqp19808:\*  
2: Geneseqp19908:\*  
3: Geneseqp20008:\*  
4: Geneseqp20018:\*  
5: Geneseqp20028:\*  
6: Geneseqp20038:\*  
7: Geneseqp20048:\*  
8: Geneseqp20058:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	138	100.0	134	4	AAG75956 Human col
2	138	100.0	314	8	ABM81421 Tumour-as
3	138	100.0	378	6	ABP98851 Human str
4	138	100.0	401	7	ADJ69563 Human hea
5	138	100.0	406	5	ABP53018 Human p50
6	138	100.0	465	5	ABM58868 Breast an
7	135	97.8	183	5	ABP53019 Mouse p50
8	133	96.4	52	5	ABP53016 Cellular
9	103	74.6	23	5	ABP52966 Cellular
10	56	40.6	334	6	ADH10326 Allicococ
11	51	37.0	932	6	ABU19329 Protein e
12	51	37.0	10917	6	AAE36132 Streptom
13	50	36.2	210	3	AAE61007 Arabidops
14	50	36.2	210	3	AAE61007 Arabidops
15	50	36.2	640	6	ABU16767 Protein e
16	50	36.2	773	4	AAU35737 Helicobac
17	50	36.2	773	5	ABU30798 Protein e
18	50	36.2	855	5	AAU99662 Human his
19	50	36.2	855	7	ADP21699 Human his
20	50	36.2	855	7	ADP21699 Human his
21	50	36.2	855	7	ADP21699 Human his
22	50	36.2	912	8	ADH43346 Human gla
23	50	36.2	915	6	ABP56824 Human his
24	50	36.2	933	7	ABM85826 Mouse pro
25	50	36.2	938	8	ADJ77693 Murine hi
25	50	36.2	1008	4	AAH78891 Human pro

26	50	36.2	1020	4	AAH79875 Human pro
27	50	36.2	1642	8	ADQ39513 Human myo
28	50	36.2	1642	8	ADQ39515 Human myo
29	50	36.2	2000	8	ADN04531 Antipsocti
30	50	36.2	2000	8	ADP23738 PRO polyp
31	50	36.2	3024	7	ADJ70171 Human hea
32	50	36.2	3396	7	ADD48597 Human pro
33	50	36.2	3396	7	ADN95526 Human BEC
34	50	36.2	3396	8	ADJ75521 Marker ge
35	50	36.2	3396	8	ADQ39509 Human myo
36	50	36.2	3396	8	ADQ39510 Human myo
37	50	36.2	3396	8	ADP99164 Chondroit
38	49	35.5	87	4	ABM38883 Peptide #
39	49	35.5	87	4	AAH32468 Human Don
40	49	35.5	87	4	AAH72208 Human bra
41	49	35.5	87	4	AAH59635 Human liv
42	49	35.5	87	4	ABG53895 Human pep
43	49	35.5	87	5	ABG42023 Human pep
44	49	35.5	279	6	ABU29933 Protein e
45	49	35.5	340	8	ADQ19111 Human sol

## ALIGNMENTS

RESULT 1  
AAH75956 standard; protein; 134 AA.  
ID AAG75956; (first entry)  
XX 03-SEP-2001  
XX Human colon cancer antigen protein SEQ ID NO:6720.  
XX Human, colon cancer; colon cancer antigen; diagnosis; detection;  
XX colorectal carcinoma; chromosome 12.  
XX Homo sapiens.  
XX MO200122920-A2.  
XX 05-APR-2001.  
XX 28-SEP-2000; 2000WO-US026524.  
XX 29-SEP-1999; 99US-0157137P.  
XX 03-NOV-1999; 99US-0163280P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Ruben SM, Barash SC, Birse CE, Rosen CA;  
XX WPI, 2001-235357/24.  
XX N-PSDB; AAH35361.  
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
XX useful for preventing, diagnosing and/or treating colorectal cancers.  
XX Claim 11, Page 8181, 9803pp; English.  
XX AAH32943 to AAH37195 and AAG7514 to AAG77788 represent human colon  
XX cancer-associated nucleic acid molecules (N) and proteins (P), where the  
XX proteins are collectively known as colon cancer antigens. The colon  
XX cancer antigens have cytostatic activity and can be used in gene therapy  
XX and vaccine production. N and P may be used in the prevention, diagnosis  
XX and treatment of diseases associated with inappropriate P expression. For  
XX example, N and P may be used to treat disorders associated with decreased  
XX expression by rectifying mutations or deletions in a patient's genome  
XX that affect the activity of P by expressing inactive proteins or to  
XX supplement the patients own production of P. Additionally, N may be used  
XX to produce the colon cancer-associated Ps, by inserting the nucleic acids  
XX into a host cell and culturing the cell to express the proteins. N and P

CC can be used in the prevention, diagnosis and treatment of colorectal  
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent  
CC sequences used in the exemplification of the present invention. N.B.  
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at  
CC time of publication, meaning no sequences are present for SEQ ID NO:1027  
CC to 1052, 7921 and 7922  
CC  
XX  
SQ Sequence 134 AA;  
Query Match 100.0%; Score 138; DB 4; Length 134;  
Best Local Similarity 100.0%; Pred. No. 1.2e-11;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 EVEKIKTVKESATEBEKLTPLYLAKQALAL 30  
Db 19 EVEKIKTVKESATEBEKLTPLYLAKQALAL 48  
RESULT 2  
ABM81421  
ID ABM81421 standard; protein; 314 AA.  
XX  
AC ABM81421;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Tumour-associated antigenic target (TAT) polypeptide PRO82191, SEQ:3675.  
XX  
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;  
KW tumour; diagnosis; cell proliferative disorder; breast cancer;  
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;  
KW central nervous system cancer; bladder cancer; pancreatic cancer;  
KW cervical cancer; melanoma; leukaemia; hybridisation probe;  
KW chromosome identification; chromosome mapping; gene mapping;  
KW gene therapy; cytostatic.  
XX  
OS Homo sapiens.  
XX  
PN WO2004030615-A2.  
XX  
PD 15-APR-2004.  
XX  
PF 29-SEP-2003; 2003WO-US028547.  
XX  
PR 02-OCT-2002; 2002US-0414971P.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Wu TD, Zhang Z, Zhou Y;  
XX  
XX WPI; 2004-347921/32.  
DR N-PSDB; ACN39497.  
XX  
XX New tumour-associated antigenic target polypeptides and nucleic acids,  
PT useful in preparing a medicament for treating or detecting a  
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or  
PT prostate cancer or tumor.  
XX  
XX  
PS Claim 12; SEQ ID NO 3675; 7273pp; English.  
XX  
XX The invention relates to human tumour-associated antigenic target (TAT)  
CC polypeptides, and their related nucleic acids. The TAT polypeptides are  
CC overexpressed in cancer tissues compared to normal tissues, and may thus  
CC serve as effective targets for the diagnosis and treatment of cancer in  
CC mammals. The invention also relates to nucleic acid and polypeptide  
CC sequences at least 80% identical to the TAT nucleic acids and  
CC polypeptides; expression vectors and host cells comprising a TAT nucleic  
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic  
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a  
CC TAT polypeptide; and methods and compositions for the treatment or  
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,  
CC antibodies, antagonists, binding molecules and compositions are useful  
CC for diagnosing or treating a cell proliferative disorder associated with

CC increased TAT expression, particularly cancers such as breast cancer,  
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder  
CC cancer, pancreatic cancer, cervical cancer, cancers of the central  
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be  
CC used as hybridisation probes, in chromosome and gene mapping, in  
CC chromosome identification and in gene therapy. The present sequence  
CC represents a TAT polypeptide of the invention  
CC  
XX  
SQ Sequence 314 AA;  
Query Match 100.0%; Score 138; DB 8; Length 314;  
Best Local Similarity 100.0%; Pred. No. 3.2e-11;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 EVEKIKTVKESATEBEKLTPLYLAKQALAL 30  
Db 29 EVEKIKTVKESATEBEKLTPLYLAKQALAL 58  
RESULT 3  
ABP98851  
ID ABP98851 standard; protein; 378 AA.  
XX  
AC ABP98851;  
XX  
DT 15-JUL-2003 (first entry)  
XX  
DE Human structural and cytoskeletal associated protein #42.  
XX  
KW Cytostatic; antiarteriosclerotic; anticonvulsant; nootropic; antianginal;  
KW neuroprotective; cerebroprotective; hypotensive; cardiant; osteopathic;  
KW antiinflammatory; antiarthritic; virocidic; gene therapy; human; stroke;  
KW structural and cytoskeleton-associated protein; SCAP; cancer; angina;  
KW atherosclerosis; epilepsy; Huntington's disease; hypertension;  
KW heart failure; osteoporosis; osteoarthritis.  
XX  
OS Homo sapiens.  
XX  
PN WO2003031940-A2.  
XX  
PD 17-APR-2003.  
XX  
PF 10-OCT-2002; 2002WO-US032851.  
XX  
PR 12-OCT-2001; 2001US-0328931P.  
XX  
PR 19-OCT-2001; 2001US-0360641P.  
XX  
PR 02-NOV-2001; 2001US-0343896P.  
XX  
PR 09-NOV-2001; 2001US-0346308P.  
XX  
PR 16-NOV-2001; 2001US-0332385P.  
XX  
PR 07-DEC-2001; 2001US-0340776P.  
XX  
PR 11-JAN-2002; 2002US-0347703P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
XX Becha SD, Bhatia U, Blake JF, Borowsky ML, Burrill JD, Chang H,  
XX Chwala NK, Elliott VS, Emerling BM, Forsythe JD, Goryad AE,  
PI Griffin JA, Hafalia AYA, Ho A, Ison CH, Kable AE, Khare R, Lal PG,  
PI Lee S, Lee EA, Lee SY, Lehr-Mason PM, Li JX, Lindquist BA, Luo W,  
PI Marquis JP, Rankumar J, Richardson TW, Sprague WW, Swarnakar A,  
PI Tang YT, Warren BA, Yang J, Yue H, Zebardjian Y, Zheng W;  
XX  
XX WPI; 2003-403125/38.  
DR N-PSDB; ACC44338.  
XX  
XX New human structural and cytoskeleton-associated proteins (SCAP) useful  
PT for diagnosing, treating and preventing diseases or conditions associated  
PT with the aberrant SCAP expression e.g. cancer, osteoporosis, or epilepsy.  
XX  
XX  
PS Claim 1; Page 298; 361pp; English.  
XX  
XX This sequence represents a novel isolated human structural and  
CC cytoskeleton-associated protein (SCAP) polypeptide. The polypeptides and  
CC polynucleotides encoding them are useful in diagnosing, treating and

preventing diseases or conditions associated with the decreased expression or over expression of SCAP, such as cell proliferative (e.g. cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's disease, stroke), heart (e.g. hypertension, heart failure, angina) and skeletal muscle disorders (e.g. osteoporosis, osteoarthritis) or viral infections. These are also useful in assessing the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of SCAP. The SCAP or its fragments are useful in screening compounds for effectiveness as agonist or antagonist of the polypeptides, or in altering the expression of the target polynucleotide and compounds that specifically bind to or modulate the activity of the polypeptide. The microarray is useful in monitoring or measuring protein-protein interactions, drug-target interactions, and gene expression profiles

Query Match	100.0%	Score 138	DB 6	Length 378
Best Local Similarity	100.0%	Pred. No. 3.9e-11		
Matches 30	Conservative 0	Mismatches 0	Indels 0	Gaps 0

**QY**      1 EVEKIKTTVESATEEKLTPVLAKOIAL 30  
         |||||  
**D6**      93 EVEKIKTTVESATEEKLTPVLAKOIAL 122  
         |||||

**RESULT 4**  
**ADJ69563**

ID	ADJ69563	standard; protein; 401 AA.
XX		
AC	ADJ69563;	
XX		

DT	06-MAY-2004 (first entry)
XX	
XX	
DB	Human heat mitochondrial protein as a therapeutic target SegID1369.
XX	

KM mitochondrial; human; screening assay; diabetes mellitus;  
KM Huntington's disease; osteoarthritis;  
KM Leber's hereditary optic neuropathy; LHON;  
KM mitochondrialencephalopathy lactic acidosis and stroke; MELAS;

KM myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;  
KM neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;  
KM osteopathic; ophthalmological; cytostatic.  
XX

OS	Homo sapiens.
XX	
PN	WC200308768-A2.
XX	

PD 23-OCT-2003.  
XX  
PF 04-APR-2003; 2003WO-US010870.  
XX

PR 12-APR-2002; 2002US-0372843P.  
PR 17-JUN-2002; 2002US-038987P.  
PR 20-SEP-2002; 2002US-0412418P.  
XX

PA (MITO-) MITOKOR.  
PA (BUCK-) BUCK INST AGE RES.  
XI  
Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM,

PI Warnock DE;  
XX  
DR WPI, 2003-845369/78.  
XX

PT Identifying a mitochondrial target for drug screening assays and for  
PT treating diseases associated with altered mitochondrial function,  
PT comprises detecting a modified polypeptide in a sample and correlating  
PT with the disease.

XX  
PS Claim 1; SEQ ID NO 1369; 180pp; English.  
XX  
CC This invention relates to novel mitochondrial targets that can be used

cc for therapeutic intervention in treating a disease associated with  
cc altered mitochondrial function. Specifically, it refers to a method for

identifying proteins of the human heart mitochondrial proteome that are useful for drug screening assays, as well as therapeutic targets. The present invention describes a method for identifying such proteins that can be used in the treatment of various diseases associated with altered mitochondrial function including diabetes mellitus, Huntington's disease, osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial encephalopathy lactic acidosis and stroke (MELAS), amyotonic epilepsy ragged red fibre syndrome (MERRF), or cancer. Accordingly, these compositions have neuroprotective, nootropic, antidiabetic, anticonvulsant, antiarthritis, osteopathic, ophthalmological and cytoskeletal activities. This polypeptide sequence is a human heart mitochondrial protein of the invention.

XX	
SQ	Sequence 401 AA;
Query Match	100.0%; Score 138; DB 7; Length 401;

Best Local Similarity 100.0%; Pred. No. 4.2e-11;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 116 EWEKITTVESATEEKLTPVILAKQJAL 145

RESULT 5  
ABP53018  
ID ABP53018 standard; protein, 406 AA.  
XX

AC	ABP53018;
XX	
XX	
DT	05-NOV-2002 (first entry)
XX	

DE	Human p50 amino acid sequence SEQ ID NO:53.
XX	
KW	Cellular proliferation inhibition; cytostatic; antiinflammatory; cancer;
KM	p50 inhibitor; dynactin inhibitor; gene therapy; tumour; carcinoma;

KM  
KM  
KM  
KM  
KM  
KM

sarcoma; glioblastoma; leukaemia; lymphoid malignancy; neuronal disorder;  
glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;  
glandular disorder; macrophagal disorder; epithelial disorder;  
stromal disorder; blastocoeleic disorder; angiogenic disorder;

KW immunologic disorder.  
 XX  
 OS Homo sapiens.  
 XX

```

PN      WQ200264779-A2.
XX
PD      22-AUG-2002.
XX
XX      22-AUG-2002.
XX

```

PF 21-JAN-2002; 2002WC-05001708.  
 XX  
 PR 14-FEB-2001; 2001US-00782816.  
 XX  
 /-2003 \ WWW.SILICON

PA (REGC ) UNIV CALIFORNIA.  
XX  
PI Sharp DJ, Rogers GC, Scholey JM;  
XX

XX New peptide inhibitors of p50/dynamin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT  
DR WPI; 2002-657599/70.

PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and immunologic disorders.  
XX  
PS Disclosure; Fig 1; 55pp; English.

XX The present invention describes an isolated peptide (I) comprising or  
CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
CC sequences given in ABPS2966 and ABPS2967 and can have C-terminal and N-  
CC terminal amino acid deletions and modifications and modifications and deletions

CC can be used as p50/dynamin inhibitors and in gene therapy. The peptides, nucleic acid molecules and methods from the present invention are useful for treating cancer by inhibiting cellular proliferation, such as in bladder, prostate, breast, kidney, bladder, breast, no location for malignant tumours, lymphatic, kidney, bladder, breast,

CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various 'head and  
CC neck tumours'; leukaemias and lymphoid malignancies, other disorders such  
CC as neuronal, glial, astrocytic, hypothalamic and other glandular,  
CC macropapagal, epithelial, stromal and blascoeolic disorders; and  
CC inflammatory, angiogenic and immunologic disorders. The present sequence  
CC represents human p50 which is given in the exemplification of the present  
XX invention

Sequence 406 AA;

Query Match	100.0%;	Score 138;	DB 5;	Length 406;
Best Local Similarity	100.0%;	Pred. No. 4.3e-11;		
Matches 30; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

```
QY      1 EVEKIKTTVKESATEEKLTPVLLAKQLAAL 30
        |||||
Db     121 EVEKIKTTVKESATEEKLTPVLLAKQLAAL 150
```

## RESULT 6

ID AAB58968 standard; protein; 465 AA.

AC AAB58968;

DT 27-MAR-2001 (first entry)

DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 676.

Human, breast cancer; ovarian cancer; cytostatic; immunosuppressive;  
 KM neotropic; neuroprotective; antiviral; antiallergic; hepatocytic;  
 KM antidiabetic; antiinflammatory; antitumor; vulnery; anticonvulsant;  
 KM anticholesterol; antifungal; antiparasitic; cardiant; immune disorder;  
 KM Addison's disease; allergy; autoimmune haemolytic anaemia;  
 KM autoimmune thyroiditis; diabetes mellitus; Croun's disease;  
 KM multiple sclerosis; rheumatoid arthritis; ulcerative colitis;  
 KM cardiovascular disorder; wound healing; neurological disease.

**Homo sapiens.**

PN WO200055173-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US005881.

PR 12-MAR-1999; 99US-0124270P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

WPI; 2000-611515/58.

XX 5  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840

PT polypeptides encoded by these genes, useful in the prevention, treatment

PT and diagnosis of cancer, immune disorders, cardiovascular disorders and neurological diseases.

PS Claim 11; Page 1126-1128; 1299pp; English

CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human  
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are  
CC associated with breast and ovarian cancer. Included in the invention are  
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the  
CC isolation and characterization of the DNA and protein sequences of the  
CC invention. The breast and ovarian cancer associated DNA, protein, agonist  
CC or antagonist sequences exhibit cytostatic; immunosuppressive; nocotripic;  
CC neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;  
CC antiinflammatory; antitumor; vulnerable; anticonvulsant; antibacterial;  
CC antifungal; antiparasitic and cardiant activity. The polynucleotide and

CC protein sequences are used in the diagnosis of cancer, particularly  
CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists  
CC and antagonists may also be used in the diagnosis, prevention and treatment  
CC of immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
CC cardiovascular disorders such as myocardial ischaemias; wound healing;  
CC neurological diseases such as cerebral amoxia and epilepsy; and  
CC infectious diseases

Query Match	100.0%	Score 138; DB 3;	Length 465;
Best Local Similarity	100.0%;	Pred. No. 5e-11;	
Matches 30; Conservative	0;	Mismatches 0;	Gaps 0

QY 1 EVEKIKTTVKESATEEKLTPVLLAKQQLAAL 30  
| | | | | | | | | | | | | | | | | | | | | |  
Db 180 EVEKIKTTVKESATEEKLTPVLLAKQQLAAL 209

## RESULT 7

ID ABP53019 standard; protein; 183 AA.

AC ABP53019,

DT 05-NOV-2002 (first entry)

DE	Mouse p50 amino acid sequence	SEQ ID NO:54
DE		

KW Cellular proliferation inhibition, cytostatic, antiinflammatory; cancer;  
KW p50 inhibitor; dynamitin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukaemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytic disorder; hypothalamic disorder; inflammatory;  
KW glauclular disorder; macrophagal disorder; epithelial disorder;  
KW stomal disorder; blastocoele disorder; angiogenic disorder;  
KW immunologic disorder.

Mus musculus.

PN WO200264779-A2.

PD 22-AUG-2002.

21-JAN-2002; 2002WO-US001708.

PR 14-FEB-2001; 2001US-00782816.

PA (REGC ) UNIV CALIFORNIA.

Sharp DJ, Rogers GC, Scholey JM,

WPI; 2002-657599/70.

PT New peptide inhibitors of p50/dynaminin useful for treating cancer by

PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and

F1 LITHOLOGIC UNITS:

XX

The present invention describes an isolated peptide (1) comprising or having at least 90% identity to (p1) or (p2) where (p1) and (p2) are the sequences given in ABP52966 and ABP52967 and can have C-terminal and N-terminal extensions. (1) have cytostatic and antiinflammatory activities and can be used as p50/dynactin inhibitors and in gene therapy. The peptides, nucleic acid molecules and methods from the present invention are useful for treating cancer by inhibiting cellular proliferation, such as benign or malignant tumours (renal, liver, kidney, bladder, breast, gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval, cervical, hepatic carcinomas, sarcomas, glioblastomas, and various head and neck tumours); leukemias and lymphoid malignancies, other disorders such

CC as neuronal, glial, astrocytal, hypochalamic and other glandular,  
CC macrophagal, epithelial, stromal and blastocoeleic disorders; and  
CC inflammatory, angiogenic and immunologic disorders. The present sequence  
CC represents mouse p50 which is given in the exemplification of the present  
CC invention

XX Sequence 183 AA;

Query Match 97.8%; Score 135; DB 5; Length 183;  
Best Local Similarity 96.7%; Pred. No. 4.5e-11;  
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVEKIKTVKESATEKLPVLAQKQAL 30  
Db 116 EVEKIKTVKESATEKLPVLAQKQAL 145

RESULT 8  
ABP53016  
ID ABP53016 standard; peptide; 52 AA.

XX ABP53016;  
XX  
XX 05-NOV-2002 (first entry)

DE Cellular proliferation peptide inhibitor SEQ ID NO:51.

XX Cellular proliferation inhibition; cytostatic; antiinflammatory; cancer;  
KW p50 inhibitor; dynamitin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukaemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytal disorder; hypochalamic disorder; inflammatory;  
KW glandular disorder; macrophagal disorder; epithelial disorder;  
KW stromal disorder; blastocoeleic disorder; angiogenic disorder;  
KW immunologic disorder.

XX Homo sapiens.  
OS Mus musculus.  
OS Synthetic.

XX Key Location/Qualifiers  
FT Misc-difference 44  
FT /label= 'Leu, Val  
FT /note= "Leu in humans and Val in Mus musculus"

XX WO200264779-A2.

XX 22-AUG-2002.

XX 21-JAN-2002; 2002WO-US001708.

XX 14-FEB-2001; 2001US-00782816.

XX (REGC ) UNIV CALIFORNIA.

XX Sharp DJ, Rogers GC, Scholey JW;

XX WPI; 2002-657599/70.

XX New peptide inhibitors of p50/dynamitin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.

XX Claim 2; Page 31; 55pp; English.

XX The present invention describes an isolated peptide (I) comprising or  
CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
CC terminal extensions. (I) have cytostatic and antiinflammatory activities  
CC and can be used as p50/dynamitin inhibitors and in gene therapy. The  
CC peptides, nucleic acid molecules and methods from the present invention  
CC are useful for treating cancer by inhibiting cellular proliferation, such  
CC as benign or malignant tumours (renal, liver, kidney, bladder, breast,

CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
CC neck tumours); leukaemias and lymphoid malignancies, other disorders such  
CC as neuronal, glial, astrocytal, hypochalamic and other glandular,  
CC macrophagal, epithelial, stromal and blastocoeleic disorders; and  
CC inflammatory, angiogenic and immunologic disorders. The present sequence  
CC represents a specifically claimed peptide inhibitor of cellular  
CC proliferation from the present invention

XX Sequence 52 AA;

Query Match 96.4%; Score 133; DB 5; Length 52;  
Best Local Similarity 96.7%; Pred. No. 2e-11;  
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVEKIKTVKESATEKLPVLAQKQAL 30  
Db 23 EVEKIKTVKESATEKLPVLAQKQAL 52

RESULT 9  
ABP52966  
ID ABP52966 standard; peptide; 23 AA.

XX ABP52966;  
XX  
XX 05-NOV-2002 (first entry)

DE Cellular proliferation inhibitor related peptide SEQ ID NO:1.

XX Cellular proliferation inhibition; cytostatic; antiinflammatory; cancer;  
KW p50 inhibitor; dynamitin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukaemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytal disorder; hypochalamic disorder; inflammatory;  
KW glandular disorder; macrophagal disorder; epithelial disorder;  
KW stromal disorder; blastocoeleic disorder; angiogenic disorder;  
KW immunologic disorder.

XX Homo sapiens.  
OS Mus musculus.  
OS Synthetic.

XX Key Location/Qualifiers  
FT Misc-difference 22  
FT /label= 'Leu, Val  
FT /note= "Leu in humans and Val in Mus musculus"

XX WO200264779-A2.

XX 22-AUG-2002.

XX 21-JAN-2002; 2002WO-US001708.

XX 14-FEB-2001; 2001US-00782816.

XX (REGC ) UNIV CALIFORNIA.

XX Sharp DJ, Rogers GC, Scholey JW;

XX WPI; 2002-657599/70.

XX New peptide inhibitors of p50/dynamitin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.

XX Claim 1; Page 29; 55pp; English.

XX The present invention describes an isolated peptide (I) comprising or  
CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
CC terminal extensions. (I) have cytostatic and antiinflammatory activities  
CC and can be used as p50/dynamitin inhibitors and in gene therapy. The

CC peptides, nucleic acid molecules and methods from the present invention  
CC are useful for treating cancer by inhibiting cellular proliferation, such  
CC as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulva,  
CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
CC neck tumors); leukemias and lymphoid malignancies, other disorders such  
CC as neuronal, glial, astrocytic, hypochalamic and other glandular,  
CC macrophagal, epithelial, stromal and blastocoele disorders; and  
CC inflammatory, angiogenic and immunologic disorders  
XX  
SQ Sequence 23 AA;

Query Match 74.6%; Score 103; DB 5; Length 23;  
Best Local Similarity 95.7%; Pred. No. 1.5e-07;  
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVEKIKTVKESATEKLTPLVL 23  
DB 1 EVEKIKTVKESATEKLTPLVL 23

## RESULT 10

ADBI0326  
ID ADBI0326 standard; protein; 334 AA.

AC ADBI0326;

DT 20-NOV-2003 (first entry)

DE Alloicoccus otitis antigenic protein SEQ ID NO:4734.

KW Alloicoccus otitidis; antigenic protein; immunogenic; immunisation;  
KW gene therapy; Gram-positive bacterium; infection.

OS Alloicoccus otitis.

PN WO2003048304-A2.

PD 12-JUN-2003.

PE 25-NOV-2002; 2002WO-US036123.

PR 29-NOV-2001; 2001US-0333777P.

PR 18-NOV-2002; 2002US-0426742P.

PA (AMHP) WYETH HOLDINGS CORP.

PI Fletcher LD, Mcmichael JC, Russell DP, Zagursky RJ;

DR WPI; 2003-505284/47.

DR N-PSDB; ADBI0329.

XX New Alloicoccus otitidis polynucleotides and polypeptides, useful for  
XX treating and diagnosing diseases, drug screening assays and monitoring of  
XX effects during drug clinical trials.

XX Claim 33; SEQ ID NO 4734; 1019pp; English.

XX The present invention describes an isolated polynucleotide (1) of  
XX Alloicoccus otitidis genomic DNA, which encodes an antigenic protein.  
XX Alloicoccus otitidis is a Gram-positive bacterium. Also described: (1)  
XX an isolated polypeptide that is encoded by the polynucleotide (1); (2) an  
XX expression vector comprising the novel isolated polynucleotide (1), its  
XX complement, degenerate variant or fragment; (3) a genetically engineered  
XX host cell, transfected, transformed or infected with the vector of (2);  
XX (4) an antibody specific for the polypeptide of (1); (5) an immunogenic  
XX composition comprising the polypeptide, its complement, biological  
XX equivalent or fragment, or the polynucleotide that is comprised in the  
XX expression vector; (6) a pharmaceutical composition comprising the  
XX polypeptide of (1) and a carrier; (7) a protein chip comprising an array  
XX of the polypeptides of (1), their biological equivalent or fragment; (8)  
XX immunising against Alloicoccus otitidis by administering to a host the  
XX immunogenic composition; (9) detecting and/or identifying Alloicoccus

CC otitidis in the biological sample; (10) a kit comprising a container  
CC containing the novel polynucleotide, its degenerate variant or fragment,  
CC or the antibody of (4); and (11) producing a polypeptide by culturing the  
CC genetically engineered host cell under conditions suitable to produce the  
CC polypeptide from the culture. (1) can be used in gene therapy. The  
CC polynucleotides, polypeptides, antibodies and compositions of the present  
CC invention can be used for treating and diagnosing diseases, drug  
CC screening assays and monitoring of effects during drug clinical trials.  
CC The polynucleotides are useful for expressing and detecting Alloicoccus  
CC otitidis. The present sequence represents an Alloicoccus otitidis  
CC antigen protein from the present invention.

XX  
SQ Sequence 334 AA;

Query Match 40.6%; Score 56; DB 6; Length 334;  
Best Local Similarity 34.5%; Pred. No. 17;  
Matches 10; Conservative 11; Mismatches 8; Indels 0; Gaps 0;

QY 2 VEKIKTVKESATEKLTPLVLAKQLAL 30  
DB 244 IDQIKSTIKSGSVKTKLGGILIKKDLSSL 272

## RESULT 11

ABU19329  
ID ABU19329 standard; protein; 932 AA.

AC ABU19329;

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #4856.

KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Borrelia burgdorferi.

PN WO20027183-A2.

PD 03-OCT-2002.

PE 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

PA (ELIT-) ELITRA PHARM INC.

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;

DR WPI; 2003-023926/02.

DR N-PSDB; ACA23199.

XX New antisense nucleic acids, useful for identifying proteins or screening  
XX for homologous nucleic acids required for cellular proliferation to  
XX isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 47253; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of  
XX the 623 antisense sequences given in the specification where expression  
XX of the nucleic acid inhibits proliferation of a cell. Also included are:  
XX (1) a vector comprising a promoter operably linked to the nucleic acid  
XX encoding a polypeptide whose expression is inhibited by the antisense  
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX polypeptide or its fragment whose expression is inhibited by the  
XX antisense nucleic acid; (4) an antibody capable of specifically binding  
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
XX proliferation or the activity of a gene in an operon required for

CC proliferation: (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
CC  
XX Sequence 932 AA:  
SQ  
Query Match 37.0%; Score 51; DB 6; Length 932;  
Best Local Similarity 37.9%; Pred. No. 2.9e+02;  
Matches 11; Conservative 6; Mismatches 12; Indels 0; Gaps 0;  
QY 2 VEKIKTVKESATEKLPVLAQQLAAL 30  
Db 397 IDKIDKATKEIYIKNKLTPTKILAQPEGAM 425  
RESULT 12  
AAE36132  
ID AAE36132 standard; protein; 10917 AA.  
XX  
AC AAE36132;  
XX  
DT 26-JUN-2003 (first entry)  
XX  
DE Streptomyces nodosus amhc gene encoded protein.  
XX  
KM Polyene; antibiotic; amphotericin; amph; polyketide synthase; enzyme.  
XX  
OS Streptomyces nodosus.  
XX  
PN WO200297082-A2.  
XX  
PD 05-DEC-2002.  
XX  
PF 27-MAY-2002; 2002WO-IE000071.  
XX  
PR 31-MAY-2001; 2001IE-00000527.  
XX  
PA (UYDU-) UNIV COLLEGE DUBLIN.  
XX  
PI Caffrey JP;  
XX  
DR WPI: 2003-201271/19.  
DR N-PSDB; AAD54645.  
XX  
XX Novel cytochrome P450 enzyme and nucleotides encoding the enzyme, useful  
PT for preparing amphotericin derivative or analog antibiotic agent with  
PT altered properties, in biosynthesis of polyketide other than  
XX amphotericin.  
PS Claim 6; Page 227-276; 276pp; English.  
XX  
CC The invention relates to the gene cluster encoding the polypeptides  
CC responsible for the biosynthesis of the polyene antibiotic amphotericin  
CC (amph) of Streptomyces nodosus. Polynucleotides of the invention are  
CC useful for preparing amphotericin derivatives or analogue antibiotic

CC agents with altered properties and in the biosynthesis of polyketides  
CC other than amphotericin. amphotril, amphotril or amphotril mutants are useful  
CC for producing amphotericin derivatives glycosylated with alternative  
CC sugars; amphotril or amphotril gene sequences are useful in engineered  
CC biosynthesis of peroxanmyl-16-methyl amphoteronolide B; amphotril and  
CC amphotril gene sequences are useful in the engineered biosynthesis of  
CC peroxanmyl-16-desacetoxy-16-methyl amphoteronolide B; amphotril, amphotril  
CC and amphotril gene sequences are useful for preparing polypeptides capable  
CC of addition of mycosamine to a polyketide other than amphoteronolide A or  
CC B or for preparing polypeptides for in vitro synthesis of GDP-mycosamine.  
CC The present sequence is polyketide synthase multienzyme housing extension  
CC modules 3, 4, 5, 6, 7 and 8 encoded by *S. nodosus* amhc gene  
XX  
SQ Sequence 10917 AA:  
Query Match 37.0%; Score 51; DB 6; Length 10917;  
Best Local Similarity 46.2%; Pred. No. 5.1e+03;  
Matches 12; Conservative 6; Mismatches 8; Indels 0; Gaps 0;  
QY 5 IKTVKESATEKLPVLAQQLAAL 30  
Db 1633 IRTVRRRAAQVSVTADGLAQLAGL 1658  
RESULT 13  
AAG61007  
ID AAG61007 standard; protein; 210 AA.  
XX  
AC AAG61007;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 79081.  
XX  
KM Protein identification, signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.  
XX  
XX Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
XX 05-MAR-1999; 99US-0123180P.  
XX 09-MAR-1999; 99US-0123548P.  
XX 23-MAR-1999; 99US-0125788P.  
XX 25-MAR-1999; 99US-0126264P.  
XX 29-MAR-1999; 99US-0126785P.  
XX 01-APR-1999; 99US-0127462P.  
XX 06-APR-1999; 99US-0128234P.  
XX 08-APR-1999; 99US-0128714P.  
XX 16-APR-1999; 99US-0129845P.  
XX 19-APR-1999; 99US-0130077P.  
XX 21-APR-1999; 99US-0130449P.  
XX 23-APR-1999; 99US-0130510P.  
XX 28-APR-1999; 99US-0130891P.  
XX 30-APR-1999; 99US-0131449P.  
XX 30-APR-1999; 99US-0132048P.  
XX 30-APR-1999; 99US-0132407P.  
XX 04-MAY-1999; 99US-0132484P.  
XX 05-MAY-1999; 99US-0132485P.  
XX 06-MAY-1999; 99US-0132486P.  
XX 06-MAY-1999; 99US-0132487P.  
XX 07-MAY-1999; 99US-0132863P.  
XX 11-MAY-1999; 99US-0134256P.  
XX 14-MAY-1999; 99US-0134218P.  
XX 14-MAY-1999; 99US-0134219P.  
XX 14-MAY-1999; 99US-0134221P.  
XX 14-MAY-1999; 99US-0134370P.



PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 04-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147319P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147433P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 18-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 16-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 22-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159684P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.

```
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161921P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161923P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 36.2%; Score 50; DB 3; Length 210;
Best Local Similarity 50.0%; Pred. No. 71;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Cy 1 EVEKIKTVKESATEK 18
Db 174 EPERVKTWIEEFLTEK 191

RESULT 14
AA017748
ID AA017748 standard; protein; 210 AA.
XX
AC AA017748;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 18886.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 26-APR-1999; 99US-0131448P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134370P.
PR 14-MAY-1999; 99US-0134768P.
PR 18-MAY-1999; 99US-0134941P.
PR 19-MAY-1999; 99US-0135124P.
PR 20-MAY-1999; 99US-0135353P.
PR 21-MAY-1999; 99US-0135352P.
PR 24-MAY-1999; 99US-0136021P.
PR 25-MAY-1999; 99US-0136392P.
PR 27-MAY-1999; 99US-0136392P.
```

```
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140931P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
```

```

PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147933P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155133P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0158293P.
PR 13-OCT-1999; 99US-0158294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

```

```

Query Match 36.2%; Score 50; DB 3; Length 210;
Best Local Similarity 50.0%; Pred. No. 71;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Oy 1 EVEKIKTTVESATEEKL 18
|::|::|::|::|::|
Db 174 EPERKTMIEEFLTEKTI 191

RESULT 15
ABUL6767
ID ABUL6767 standard; protein; 640 AA.
XX
XX ABUL6767;
AC
XX 19-JUN-2003 (first entry)
DT
XX
XX Protein encoded by Prokaryotic essential gene #2294.
DE
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX Acinetobacter baumannii.
OS
XX
XX WO20027183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
PF
XX
XX 21-MAR-2001; 2001US-00815242.
PR
XX 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
ER
XX
XX (ELIT-) ELITRA PHARM INC.
PA
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
PI
XX
XX WPI; 2003-029926/02.
DR
XX
XX N-PSDB; ACAA20637.
XX
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation or
PT isolate candidate molecules for rational drug discovery programs.
PT
XX
XX
XX Claim 25; SEQ ID NO 44691; 1766p; English.
XX
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required

```



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 3, 2005, 21:43:09 ; Search time 10.9091 Seconds  
(without alignments)  
264.596 Million cell updates/sec

Title: 09782816-1-22

Sequence: 1 EWEKIKTTWKESATEEKLTPVLAKQALAL 30

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 segs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR\_79:\*  
1: piri:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	38.4	1394	2 T34061	hypothetical prote
2	51	37.0	773	2 B71931	DNA gyrase chain B
3	51	37.0	932	2 C70191	penicillin-binding
4	50.5	36.6	191	2 AF0580	lipase-protein 11
5	50	36.2	210	2 B84499	hypothetical prote
6	50	36.2	773	2 B64582	DNA gyrase, sub B
7	50	36.2	878	2 B64582	hypothetical prote
8	50	36.2	1151	2 AG1717	probable peptidogl
9	49	35.5	298	2 B82707	carboxyphosphonen
10	49	35.5	340	2 T42684	hypothetical prote
11	49	35.5	455	2 B71480	hypothetical prote
12	48.5	35.1	311	1 IQBS44	probable agx-1 hom
13	48	34.8	211	2 T52431	primosome componen
14	48	34.8	379	2 T52431	hypothetical prote
15	48	34.8	1112	2 D75056	hypothetical prote
16	48	34.8	1287	2 T22235	cell division cont
17	48	34.8	2515	2 T22235	hypothetical prote
18	47.5	34.4	225	2 E84423	posterior-group pr
19	47	34.1	103	2 T07234	hypothetical prote
20	47	34.1	138	2 AD1154	ribosomal protein
21	47	34.1	138	2 AH1512	hypothetical prote
22	47	34.1	227	2 E75066	probable translati
23	47	34.1	445	2 B75558	acetyl-CoA carboxy
24	47	34.1	465	2 H87636	glutamine syntheta
25	47	34.1	697	2 T03834	nuclear distributi
26	47	34.1	761	1 S61642	translation elonga
27	46.5	33.7	453	2 S38966	hydroxymethylgluta
28	46.5	33.7	543	2 T34481	hypothetical prote
29	46.5	33.7	1964	2 T09059	notch4 - mouse

30	46	33.3	124	2 AF0088	hypothetical prote
31	46	33.3	190	2 A88961	protein F59A7.4 [i
32	46	33.3	205	2 I46876	phosphoprotein pho
33	46	33.3	482	2 S75847	signal recognition
34	46	33.3	554	2 F86244	hypothetical prote
35	46	33.3	560	2 S53382	protein YKR029c ho
36	46	33.3	574	2 T51799	CtB1-like protein
37	46	33.3	710	2 F86778	DNA topoisomerase
38	46	33.3	727	2 AD2168	hypothetical prote
39	46	33.3	757	1 S48841	secretory componen
40	46	33.3	757	1 I45956	polymeric immunogl
41	46	33.3	1103	2 JC4114	Ca2+-transporting
42	46	33.3	1612	2 AB1347	probable peptidogl
43	46	33.3	2254	2 D86215	protein T6D22.14 [
44	45.5	33.0	253	2 T37247	histone H1.4 - Cae
45	45.5	33.0	253	2 F89030	protein C18G1.5 [i

## ALIGNMENTS

## RESULT 1

T34061 hypothetical protein F28B3.1 - Caenorhabditis elegans

CISpecies: Caenorhabditis elegans  
CDate: 29-Oct-1999 #sequence revision 29-Oct-1999 #text\_change 09-Jul-2004

CAccession: T34061  
RGeisel, C.; Kramer, J.; Smith, A.

A:Submitted to the EMBL Data Library, May 1997  
A:Description: The sequence of C. elegans cosmid F28B3.

A:Reference number: Z21469  
A:Accession: T34061

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA

A:Residues: 1-1394 <GBI>  
A:Cross-references: UNIPROT:O01787; EMBL:AF003136; PIDD:AA93633.1; GSPDB:GN00019; CESP:

A:Experimental source: strain Bristol N2; clone F28B3  
C:Genetics:

A:Gene: CESP:F28B3.1  
A:Map position: 1

A:Introns: 46/3; 105/2; 280/3; 337/1; 358/1; 530/3; 685/3; 789/3; 1311/2

Query Match	38.4%	Score 53;	DB 2;	Length 1394;
Best Local Similarity	40.9%	Pred. No. 67;		
Matches	9;	Conservative	7;	Mismatches 6;
				Indels 0;
				Gaps 0;
Cy	3	EKIKTTWKESATEEKLTPVLA 24		
Db	1370	EKLSTTRSLFLEKLNPIYVA 1391		

## RESULT 2

B71931 DNA gyrase chain B - Helicobacter pylori (strain J99)

CISpecies: Helicobacter pylori  
A:Title: stralin J99  
A:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004

CAccession: B71931  
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999  
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A:Reference number: A71800; MUID:99120557; PMID:9923682  
A:Accession: B71931

A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-773 <ARN>  
A:Cross-references: UNIPROT:Q92LX3; GB:AE001479; GB:AE001439; NID:g4154979; PIDD:AA0603

A:Experimental source: strain J99  
C:Genetics:

A:Gene: gyrb  
C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain B

Query Match 37.0%; Score 51; DB 2; Length 773;  
Best Local Similarity 43.5%; Pred. No. 66;  
Matches 10; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

OY 5 IKTIVKESATEBKLTPLVLAQQL 27  
DB 222 LKQFVKDSAKKELTPIISFKSM 244

RESULT 3  
C70191  
penicillin-binding protein (pbp-3) homolog - Lyme disease spirochete  
C/Species: Borrelia burgdorferi (Lyme disease spirochete)  
C/Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 09-Jul-2004  
C/Accession: C70191  
R/Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White  
son, D.; Peterson, J.; Kellavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vagt,  
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
Nature 390, 580-586, 1997  
A/Authors: Smith, H.O.; Venter, J.C.  
A/Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.  
A/Reference number: A70100; MUID:98065943; PMID:9403685  
A/Accession: C70191  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-932 <KLE>  
A/Cross-references: UNIPROT:O51674; GB:AE001173; GB:AE000783; NID:g2688665; PIDN:AAC6708  
A/Experimental source: strain B31

Query Match 37.0%; Score 51; DB 2; Length 932;  
Best Local Similarity 37.9%; Pred. No. 81;  
Matches 11; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

OY 2 VEKIKTVKESATEBKLTPLVLAQQL 30  
DB 397 IDKIDKATKEITIKKLTPLVLAQPEGAM 425

RESULT 4  
AF0580  
lipoate-protein ligase B (lipoate biosynthesis protein B) [imported] - Salmonella enteri  
C/Species: Salmonella enterica subsp. enterica serovar Typh  
A/Note: this species has also been called Salmonella typhi  
C/Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C/Accession: AF0580  
R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
Th, T.; Comercon, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
S.; Koulle, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica sero  
A/Reference number: AB0502; MUID:21534947; PMID:11677608  
A/Accession: AF0580  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-191 <PAR>  
A/Cross-references: GB:AL513382; PIDN:CAD05112.1; PID:g16501886; GSPDB:GN00176  
C/Genetics:  
A/Gene: STY0686  
C/Superfamily: Escherichia coli lipoate-protein ligase lipB

Query Match 36.6%; Score 50.5; DB 2; Length 191;  
Best Local Similarity 50.0%; Pred. No. 17;  
Matches 15; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

OY 1 EVEKIKTVKESATEBKLTPLVLAQQL 30  
DB 153 EMAKI-TQWKSDATDNIAPRLNITLAL 181

RESULT 5  
B84499  
hypothetical protein At2g11890 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 12-Jul-2004  
C/Accession: B84499  
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Unayam, L.; Talion, L.;  
Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J  
Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: AB4420; MUID:20083487; PMID:10617197  
A/Accession: B84499  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-210 <STO>  
A/Cross-references: UNIPROT:Q9S1Y3; GB:AE002093; NID:g4557062; PIDN:AAD22501.1; GSPDB:GN  
C/Genetics:  
A/Gene: At2g11890  
A/Map position: 2  
C/Superfamily: uncharacterized CYTH domain protein

Query Match 36.2%; Score 50; DB 2; Length 210;  
Best Local Similarity 50.0%; Pred. No. 22;  
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 EVEKIKTVKESATEBKLT 18  
DB 174 EPERVKTMIEEFLTEKTI 191

RESULT 6  
B64582  
DNA gyrase, sub B - Helicobacter pylori (strain 26695)  
C/Species: Helicobacter pylori  
C/Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004  
C/Accession: B64582  
R/Tomb, J.F.; White, O.; Kellavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wathey, L.  
Nature 388, 539-547, 1997  
A/Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser, C.  
A/Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A/Reference number: A64520; MUID:97394467; PMID:9252185  
A/Accession: B64582  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-773 <TOM>  
A/Cross-references: UNIPROT:P55992; GB:AE000564; GB:AE000511; NID:g2313602; PIDN:AAD0756  
C/Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain B

Query Match 36.2%; Score 50; DB 2; Length 773;  
Best Local Similarity 43.5%; Pred. No. 90;  
Matches 10; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

OY 5 IKTIVKESATEBKLTPLVLAQQL 27  
DB 222 LKQFVKDSAKKELTPIISFKSM 244

RESULT 7  
T17245  
hypothetical protein DKFZp586J0917.1 - human (fragment)  
C/Species: Homo sapiens (man)  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C/Accession: T17245  
R/Kocher, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, September 1999  
A/Reference number: Z18722  
A/Accession: T17245  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-878 <KOB>  
A/Cross-references: EMBL:AL117455  
A/Experimental source: adult uterus; clone DKFZp586J0917  
C/Genetics:



A:Note: DKFZp586J0917.1

Query Match

Best Local Similarity 36.2%; Score 50; DB 2; Length 878;  
Matches 11; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 6 KTTVESATEEKLTPVLAKQAL 30

DB 3 RSAVSSVVKQKLAELILKQQAAL 27

RESULT 8

AG1717

probable peptidoglycan bound protein (LPXTG motif) lin2283 [imported] - *Listeria innocua*  
C:Species: *Listeria innocua*  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004

C:Accession: AG1717

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, O., C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AG1717

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1151 <GAA>  
A:Cross-references: UNIPROT:Q929J2; GB:AE0592022; PIDN:CA097511.1; PID:gl6414795; GSPDB:C

A:Experimental source: strain C1p11262  
C:Genetics:  
A:Gene: lin2283

Query Match

Best Local Similarity 36.2%; Score 50; DB 2; Length 1151;  
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 EKIKTVESATEEKL 19

DB 31 EGVETNVEATTEKVT 47

RESULT 9

B82707  
carboxyphosphonoenolpyruvate phosphonmutase XF1234 [imported] - *Xylella fastidiosa* (str

C:Species: *Xylella fastidiosa*  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004

C:Accession: B82707  
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.

A:Reference number: A82515; MUID:20365717; PMID:10910347  
A:Note: for a complete list of authors see reference number AS9328 below

A:Accession: B82707

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-298 <SIM>  
A:Cross-references: UNIPROT:Q9P24; GB:AE003957; GB:AE003849; NID:9106207; PIDN:AA08404

A:Experimental source: strain 945C  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre, H

as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fromt

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, E

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: AS9328

A:Contents: annotation

C:Genetics:

A:Gene: XF1234  
C:Superfamily: carboxyphosphonoenolpyruvate phosphonmutase

Query Match

Best Local Similarity 35.5%; Score 49; DB 2; Length 298;  
Matches 12; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 2 VEKIKTVESATEEKLTPVLAKQAL 30

DB 204 VERTRPVPLAMTEFGKTPILTRQLEAV 232

RESULT 10

B71480  
hypothetical protein DKFZp434F232.1 - human (fragment).

C:Species: *Homo sapiens* (man)  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004

C:Accession: T42684

R:Blum, H.; Bauerachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, November 1999

A:Reference number: Z22233

A:Accession: T42684

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-340 <AAA>

A:Cross-references: UNIPROT:Q96MK3; EMBL:AL133105

A:Experimental source: adult testis; clone DKFZp434F2322

C:Genetics:  
A:Note: DKFZp434F2322.1

Query Match

Best Local Similarity 35.5%; Score 49; DB 2; Length 340;  
Matches 10; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 4 KIKTVESATEEKLTPVLAKQAL 30

DB 271 RLSDVMRESLLELDQSLPVLTEPHLAL 297

RESULT 11

B71480  
probable agx-1 homolog- $\alpha$ -glucose pyrophosphorylase - *Chlamydia trachomatis* (serotype L

C:Species: *Chlamydia trachomatis*  
C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 09-Jul-2004

C:Accession: B71480  
R:Stephens, R.S.; Kalman, S.; Lammel, C.D.; Fan, J.; Marache, R.; Aravind, L.; Mitchell,

Science 282, 754-759, 1998  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trac*

A:Reference number: A71570; MUID:9900809; PMID:9784136  
A:Accession: B71480

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-455 <ARN>

A:Cross-references: UNIPROT:O84720; GB:AE001342; GB:AE001273; NID:93329166; PIDN:AA06831

A:Experimental source: serotype D, strain UW-3/Cx  
C:Genetics:  
A:Gene: CT715

Query Match  
Best Local Similarity 35.5%; Score 49; DB 2; Length 455;  
Matches 12; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 5 IKTVESATEEKLTPVLAKQAL 30

DB 264 IKTLRQSAQEDVGLILAKQIAV 289

RESULT 12

IOB844  
primosome component (helicase loader) dnaI - *Bacillus subtilis*  
N:Alternate names: dnaI protein homolog, 44K, hypothetical protein Y (dnaB 3' region)

C/Species: Bacillus subtilis  
C/Date: 31-Mar-1993 #sequence revision 31-Mar-1993 #text\_change 09-Jul-2004  
C/Accession: B24720; C26580; F69617  
R/Ogasawara, N.; Moriya, S.; Maza, P.G.; Yoshikawa, H.  
Nucleic Acids Res. 14, 9989-9999, 1986  
A/Title: Nucleotide sequence and organization of dnaB gene and neighbouring genes on the  
A/Reference number: A93650; MUID:87117549; PMID:3027671  
A/Accession: B24720  
A/Molecule type: DNA  
A/Residues: 1-311 <OGA>  
A/Cross-references: UNIPROT:P06567; GB:X04963; NID:939880; PIDN:CAA28633.1; PID:939881  
R/Hoshino, T.; McEnzie, T.; Schmitz, S.; Tanaka, T.; Sueoka, N.  
Proc. Natl. Acad. Sci. U.S.A. 84, 653-657, 1987  
A/Title: Nucleotide sequence of Bacillus subtilis dnaB: a gene essential for DNA replica  
A/Reference number: A94709; MUID:87118226; PMID:3027697  
A/Accession: C26580  
A/Molecule type: DNA  
A/Residues: 1-18, 'N', 20-23, 'T', 25-206 <HOS>  
R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter  
C.; Bron, S.; Brouillet, S.; Bursch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chc  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, B.  
Nature 390, 249-256, 1997  
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall  
leoch, U.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.  
Koeber, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A.; Autors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, V.; Masuda, S.; Maueel  
Y., M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetall  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadle, Y.; Sato, T.; Scanlon,  
A.; Autors: Schleich, S.; Schroeder, R.; Scoffone, F.; Sekiguchi, J.; Sekowka, A.; Serz  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K  
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A/Reference number: A69580; MUID:98044033; PMID:9384377  
A/Accession: F69617  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-311 <KUN>  
A/Cross-references: GB:293918; GB:AL009126; NID:92635200; PIDN:CAB14858.1; PID:92635363  
A/Experimental source: strain 168  
C/Genetics:  
A/Gene: dnaI  
C/Superfamily: 44K dnaA protein homolog  
C/Keywords: ATP; nucleotide binding; P-loop  
F:168-175/Region: nucleotide-binding motif A (P-loop)  
Query Match 35.1%; Score 48.5; DB 1; Length 311;  
Best Local Similarity 33.3%; Pred. No. 54;  
Matches 11; Conservative 8; Mismatches 9; Indels 5; Gaps 1;  
Oy 2 VEKIKTVKESATEKLT-----TPVLAKQLAA 29  
Db 201 VREIKNSLDQTLERKLMNVKTTPLVLMDDIGA 233  
RESULT 13  
T52241  
hypothetical protein zcf37 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 24-Oct-2000 #sequence\_revision 24-Oct-2000 #text\_change 09-Jul-2004  
C/Accession: T52431  
R/Kao, A.; Suzuki, M.; Kuwahara, A.; Ooe, H.; Higano-Inaba, K.; Komeda, Y.  
Gene 239, 309-316, 1999  
A/Title: Isolation and analysis of cDNA within a 300 kb Arabidopsis thaliana genomic reg  
A/Reference number: Z25171  
A/Accession: T52431  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-211 <KAT>  
A/Cross-references: UNIPROT:Q9SL79; EMBL:AB028223; PIDN:BA87954.1  
A/Experimental source: cultivar Columbia  
C/Genetics:  
A/Map position: 1

A/Note: ZCF37  
Query Match 34.8%; Score 48; DB 2; Length 211;  
Best Local Similarity 64.3%; Pred. No. 42;  
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
Oy 3 EKIKTVEKESATEE 16  
Db 91 DKVTEVKEATEE 104  
RESULT 14  
T34100  
hypothetical protein C17G10.2 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T34100  
R/Johnson, D.  
submitted to the EMBL Data Library, June 1995  
A/Description: The sequence of C. elegans cosmid C17G10.  
A/Reference number: Z21476  
A/Accession: T34100  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-379 <JOH>  
A/Cross-references: UNIPROT:Q09974; EMBL:U28739; PIDN:AAB93452.1; GSPDB:GN00020; CESP:CI  
A/Experimental source: strain Bristol N2; clone C17G10  
C/Genetics:  
A/Gene: CESP:C17G10.2  
A/Map position: 2  
A/Intons: 36/3; 81/1; 167/1; 252/3; 355/3  
Query Match 34.8%; Score 48; DB 2; Length 379;  
Best Local Similarity 32.1%; Pred. No. 79;  
Matches 9; Conservative 9; Mismatches 10; Indels 0; Gaps 0;  
Oy 2 VEKIKTVKESATEKLTTPVLAKQLAA 29  
Db 75 IEALQASIKENSPPRKLMNAVLYFRANA 102  
RESULT 15  
D75056  
cell division control protein. PAB2373 - Pyrococcus abyssi (strain Orsay)  
C/Species: Pyrococcus abyssi  
C/Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C/Accession: D75056  
R/anonymous, Genoscope  
submitted to the EMBL Data Library, July 1999  
A/Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome struc  
A/Reference number: A75001  
A/Accession: D75056  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-1112 <RAW>  
A/Cross-references: UNIPROT:Q09YR7; GB:A242827; GB:AL096836; NID:95458657; PIDN:CAB5034  
A/Experimental source: strain Orsay  
C/Genetics:  
A/Gene: cdc21; PAB2373  
Query Match 34.8%; Score 48; DB 2; Length 1112;  
Best Local Similarity 40.0%; Pred. No. 2.5e+02;  
Matches 14; Conservative 11; Mismatches 4; Indels 6; Gaps 3;  
Oy 1 EVE---KIKTVKESATEKLTPLV-LIAQLAL 30  
Db 958 EIEKYVYMKRSYKKS-SEERIKPIPTARQLAL 991  
Search completed: November 3, 2005, 22:04:14  
Job time : 11.9091 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2005, 21:37:29 ; Search time 91.4866 Seconds  
(without alignments)  
291.060 Million cell updates/sec

Title: 09782816-51

Sequence: 1 GVKETPOOKYORLHEVQEL.....ESATBEKLPVLAKQIALAL 52

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_tramb1.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	251	100.0	400	1 DCT2_HUMAN	Q13561 homo sapien
2	248	98.8	401	1 DCT2_MOUSE	Q93KJ8 mus musculu
3	248	98.8	402	1 O6A7H5	Q6A7H5 ratius notv
4	220	87.6	403	2 O61R33	Q61R33 xenopus lae
5	213	84.9	403	2 O66J30	Q66J30 xenopus lae
6	211	84.1	338	2 O72XY2	Q72XY2 xenopus lae
7	177	70.5	402	2 O9PTG6	Q9PTG6 gallus gall
8	177	70.5	405	2 O7T3H1	Q7T3H1 brachydanio
9	76	30.3	402	2 O61P53	Q61P53 xenopus lae
10	72.5	28.9	380	2 O9V4Y9	Q9V4Y9 drosophila
11	71	28.3	800	2 O6MVP7	Q6MVP7 neurospora
12	71	28.3	813	2 O7SH14	Q7SH14 neurospora
13	70	27.9	311	2 O971V5	Q971V5 sulfolobus
14	70	27.9	751	2 O8M0K1	Q8M0K1 drosophila
15	70	27.9	1087	2 O8B8S5	Q8B8S5 drosophila
16	70	27.9	1124	2 O8M0J8	Q8M0J8 drosophila
17	70	27.9	2199	2 O7PUP2	Q7PUP2 anopheles g
18	69.5	27.7	873	2 O95X56	Q95X56 caenorhabdi
19	69.5	27.7	919	2 O95X55	Q95X55 caenorhabdi
20	68	27.1	1755	2 O7RV31	Q7RV31 neurospora
21	68	27.1	1968	2 O8X0C5	Q8X0C5 neurospora
22	67.5	26.9	1868	2 O6BNV2	Q6BNV2 debrayomyce
23	67	26.7	455	2 O7ZVFL	Q7ZVFL brachydanio
24	67	26.7	639	2 O6Z245	Q6Z245 caenorhabdi
25	67	26.7	1795	2 O9LCT9	Q9LCT9 staphylococ
26	67	26.7	2478	2 O9LCT9	Q9LCT9 staphylococ
27	67	26.7	2478	2 O9LCT9	Q9LCT9 staphylococ
28	67	26.7	2481	2 O99OR6	Q99OR6 staphylococ
29	67	26.7	2481	2 O7A4B1	Q7A4B1 staphylococ
30	66	26.3	393	2 O49567	Q49567 arabidopsis
31	66	26.3	860	2 O7Q0Q9	Q7Q0Q9 anopheles g

32	66	26.3	866	2 O84500	O84500 chlamydia t
33	66	26.3	1022	2 O8TBV8	O8TBV8 homo sapien
34	65.5	26.1	1837	2 O74424	O74424 schizosach
35	65	25.9	388	2 O7P225	O7P225 anopheles g
36	64.5	25.7	348	2 O8LIJ9	O8LIJ9 oryza sativ
37	64	25.5	149	2 O9BZS3	O9BZS3 homo sapien
38	64	25.5	188	2 O9UM00	O9UM00 homo sapien
39	64	25.5	188	2 O921L3	O921L3 mus musculu
40	64	25.5	188	2 O6DGM9	O6DGM9 brachydanio
41	64	25.5	210	2 O981Y3	O981Y3 arabidopsis
42	64	25.5	230	2 O75545	O75545 homo sapien
43	64	25.5	1227	1 JIP3_DROME	JIP3_DROME
44	64	25.5	1956	2 O9Y2K3	O9Y2K3 homo sapien
45	63.5	25.3	163	1 Y012_BPHF1	Y012_BPHF1 bacterioph

## ALIGNMENTS

RESULT 1  
DCT2\_HUMAN STANDARD; PRT; 400 AA.  
ID Q13561; O86YN2; Q9BW17;  
AC 01-NOV-1997 (Rel. 35, Created)  
DT 05-01-2004 (Rel. 44, Last sequence update)  
DT 05-01-2004 (Rel. 44, Last annotation update)  
DE Dynactin complex 50 kDa subunit (50 kDa dynein-associated polypeptide)  
DE (p50 dynactin) (DCTN-50) (Dynactin 2).  
OS Name=DCTN2; Synonyms=DCTN50;  
GN Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=9617072; PubMed=8647893; DOI=10.1083/jcb.132.4.617;  
RA Echeverri C.J., Paschal B.M., Vaughan K.T., Vallee R.B.;  
RT "Molecular characterization of the 50-kD subunit of dynactin reveals  
RT function for the complex in chromosome alignment and spindle  
RT organization during mitosis."  
RT J. Cell Biol. 132:617-633(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta, Skin, and Uterus;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Stuenkel R.D., Collins F.S., Wagner C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,  
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stempleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Stromstedt M.J., Usdin T.B., Toshiyuki S., Carrincci P., Prange C.,  
RA Baha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Roach S.A., McEwan P.J., McKernan K.J., Malek J.A., Gamaralle P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,  
RA Vallayon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fanning J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeshel R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
RA Scherch A., Schein J.E., Jones S.J.W., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [3]  
RP SEQUENCE OF 1-384 FROM N.A.  
RA Almeida J.P., Yu-lee L.-Y.;  
RT "Human 50 kD dynactin subunit, p50 dynactin, isolated from HeLa  
RT cells."  
RL Submitted (NOV-2002) to the EMBL/Genbank/DBJ databases.  
RN [4]  
RP SEQUENCE OF 1-13.



OY 1 GVKETPOOKYORLHEVOELTTEVEKITYKESATEKRLPVYLAQKLAAL 52  
DB 93 GVKETPOOKYORLHEVOELTTEVEKITYKESATEKRLPVYLAQKLAAL 144

RESULT 3  
ID 06AYH5 PRELIMINARY; PRT; 402 AA.  
AC 06AYH5.  
DT 25-OCT-2004 (TREMBlrel. 28, Created)  
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)  
DE Dynactin 2.  
GN Name=Dctn2;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
NCBI\_TaxID=10116;  
[1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Testis;  
RC PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,  
RA Jones S.J., Maitra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RA Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC079042; AAH79042.1; -  
DR GO: GO:0005869; C:dynactin complex; IEA.  
DR GO: GO:0007017; P:microtubule-based process; IEA.  
DR InterPro: IPR006996; Dynamitin.  
DR Pfam: PF04912; Dynamitin; 1.  
SQ SEQUENCE 402 AA; 44148 MW; 550335535A4FB052 CRC64;

Query Match 98.8%; Score 248; DB 2; Length 402;  
Best Local Similarity 98.1%; Pred. No. 9.4e-17;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GVKETPOOKYORLHEVOELTTEVEKITYKESATEKRLPVYLAQKLAAL 52  
DB 94 GVKETPOOKYORLHEVOELTTEVEKITYKESATEKRLPVYLAQKLAAL 145

RESULT 4  
ID 06IRB3 PRELIMINARY; PRT; 403 AA.  
AC 06IRB3.  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE Dctn2-prov protein.  
GN Name=dctn2-prov;  
OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
NCBI\_TaxID=83355;  
[1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Embryo;  
RC MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,  
RA Jones S.J., Maitra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RA MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.;  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT initiative.";  
RL Dev. Dyn. 225:384-391 (2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RA Klein S., Strausberg R.;  
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC070987; AAH70987.1; -  
DR GO: GO:0005869; C:dynactin complex; IEA.  
DR GO: GO:0007017; P:microtubule-based process; IEA.  
DR InterPro: IPR006996; Dynamitin.  
DR Pfam: PF04912; Dynamitin; 1.  
SQ SEQUENCE 403 AA; 44828 MW; 9FD158C7C7983062 CRC64;

Query Match 87.6%; Score 220; DB 2; Length 403;  
Best Local Similarity 84.6%; Pred. No. 6e-14;  
Matches 44; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 1 GVKETPOOKYORLHEVOELTTEVEKITYKESATEKRLPVYLAQKLAAL 52  
DB 94 GVKETPOOKYORLHEVOELTTEVEKITYKESATEKRLPVYLAQKLAAL 145

RESULT 5  
ID 06GJ30 PRELIMINARY; PRT; 403 AA.  
AC 06GJ30.  
DT 25-OCT-2004 (TREMBlrel. 28, Created)  
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)  
DE MGC82128 protein.  
GN Name=MGC82128;  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
NCBI\_TaxID=83355;  
[1]

RA SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.,  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT initiative.";  
RL Dev. Dyn. 225:384-391(2002).  
[2]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Kidney;  
RC PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.F., Bhat N.K.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carrinchi P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalek U., Smallos D.E., Scherch A., Schein J.E.,  
RA Jones S.J., Maitra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[3]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Kidney;  
RA Klein S., Gerthard D.S.,  
RL Submitted (Aug-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC081081; AAH81081.1;  
DR InterPro: IPR006996; Dymatitin.  
DR Pfam: PF04912; Dymatitin; 1.  
SQ SEQUENCE 403 AA; 44737 MW; 86B8CBE54325EF3 CRC64;  
Query Match 84.9%; Score 213; DB 2; Length 403;  
Best Local Similarity 82.7%; Pred. No. 3e-13;  
Matches 43; Conservative 5; Mismatches 4; Indels 0; Gaps 0;  
OY 1 GVKTPOOKYORLHVEOELTTEVEKIKTTVKESATBEKLTPLVALAKOVAL 52  
Db 94 GKKEIPPOOKYORLHVEOELTTEVEKIKTTVKESATBEKLTPLVALAKOVAL 145  
RESULT 6  
O7ZXV2 PRELIMINARY; PRT; 338 AA.  
AC O7ZXV2;  
DT 01-JUN-2003 (TREMBLrel. 24, Created)  
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Dcch2-prov protein.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.F., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carrinchi P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalek U., Smallos D.E., Scherch A., Schein J.E.,  
RA Jones S.J., Maitra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[2]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Embryo;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.,  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT initiative.";  
RL Dev. Dyn. 225:384-391(2002).  
[3]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Embryo;  
RA Klein S., Strausberg R.,  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC044069; AAH44069.1;  
DR GO; GO:0005869; C:dynactin complex; IEA.  
DR GO; GO:0007017; P:microtubule-based process; IEA.  
DR InterPro: IPR006996; Dymatitin.  
DR Pfam: PF04912; Dymatitin; 1.  
SQ SEQUENCE 338 AA; 38257 MW; 3346E0F0F644B186 CRC64;  
Query Match 84.1%; Score 211; DB 2; Length 338;  
Best Local Similarity 86.0%; Pred. No. 4e-13;  
Matches 43; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
OY 3 KETPOOKYORLHVEOELTTEVEKIKTTVKESATBEKLTPLVALAKOVAL 52  
Db 31 KETPOOKYORLHVEOELTTEVEKIKTTVKESATBEKLTPLVALAKOVAL 80  
RESULT 7  
O9PTG6 PRELIMINARY; PRT; 402 AA.  
AC O9PTG6;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Dymatitin.  
GN Name=p50;  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianidae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=20056093; PubMed=10588646;  
RA Valetti C., Metzel D.M., Schrader M., Hasbani M.J., Gill S.R.,  
RA Kreis T.E., Schroer T.A.,  
RT "Role of dynactin in endocytic traffic: effects of dymatitin  
RT overexpression and colocalization with CLIP-170.";  
RL Mol. Biol. Cell 10:4107-4120(1999).  
[2]  
RN SEQUENCE FROM N.A.  
RP Schroer T.A., Gill S.R., Hasbani M.J., Grego C.,  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF200744; AAF13996.1;  
DR GO; GO:0005869; C:dynactin complex; IEA.

DR GO:0007017; P:microtubule-based process; IEA.  
 DR InterPro; IPR006996; Dynamitin.  
 DR Pfam; PF04912; Dynamitin; 1.  
 SQ SEQUENCE 402 AA; 45126 MW; P229C467C630DCB9 CRC64;  
 Query Match 70.5%; Score 177; DB 2; Length 402;  
 Best Local Similarity 67.3%; Pred. No. 1.2e-09;  
 Matches 35; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 GVKETPOOKYORLHVEVOLTTEVEKIKTYKESATEKTLPPVLAQOLAL 52  
 DB 96 GAKETPOORVORLQHEVQELIRDVQIQSAVKESAABELTPMALAQLEBL 147

RESULT 8  
 Q773H1 PRELIMINARY; PRT; 405 AA.  
 ID Q773H1  
 AC Q773H1  
 DT 01-OCT-2003 (TRENBLrel. 25, Created)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
 DE Similar to dynactin 2 (P50).  
 GN ORENAMES=zgc:63867;  
 OS Brachydanio rerio (zebrafish) (Danio rerio)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OC NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carrincci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahney J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,  
 RA Jones S.U., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; BC053120; AAHS3120.1; -;  
 DR ZFIN; ZDB-GENE-040426-1279; zgc:63867.  
 DR GO:0005869; C:dynactin complex; IEA.  
 DR GO:0007017; P:microtubule-based process; IEA.  
 DR InterPro; IPR006996; Dynamitin.  
 DR Pfam; PF04912; Dynamitin; 1.  
 SQ SEQUENCE 405 AA; 44625 MW; CA00047342500953 CRC64;

Query Match 70.5%; Score 177; DB 2; Length 405;  
 Best Local Similarity 67.3%; Pred. No. 1.2e-09;  
 Matches 35; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 GVKETPOOKYORLHVEVOLTTEVEKIKTYKESATEKTLPPVLAQOLAL 52  
 DB 96 GVKETPOOKYORLHVEVOLTTEVEKIKTYKESATEKTLPPVLAQOLAL 147

RESULT 9  
 Q6IP53 PRELIMINARY; PRT; 402 AA.  
 ID Q6IP53  
 AC Q6IP53  
 DT 05-JUN-2004 (TRENBLrel. 27, Created)  
 DT 05-JUN-2004 (TRENBLrel. 27, Last sequence update)  
 DT 05-JUN-2004 (TRENBLrel. 27, Last annotation update)  
 DE MGC78949 protein.  
 GN Name=MGC78949;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carrincci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahney J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,  
 RA Jones S.U., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S.L., Strauberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative.";  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX Klein S., Strauberg R.;  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC072065; AAH72065.1; -;  
 DR GO:0005874; C:microtubule; IEA.  
 DR GO:0000226; P:microtubule cytoskeleton organization and b. . .; IEA.  
 DR InterPro; IPR000435; Tektin.  
 DR Pfam; PF03148; Tektin; 1.  
 DR PRINTS; PR00511; TEKTN.  
 SQ SEQUENCE 402 AA; 46535 MW; F173BC48BC0E798A CRC64;

Query Match 30.3%; Score 76; DB 2; Length 402;  
 Best Local Similarity 46.4%; Pred. No. 16;  
 Matches 13; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

QY 12 RLHVEVOLTTEVEKIKTYKESATEK 39  
 DB 336 RLHVEVOLTTEVEKIKTYKESATEK 363

RESULT 10  
 Q9V4Y9 PRELIMINARY; PRT; 380 AA.  
 ID Q9V4Y9



AC 09V4Y9;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 25-OCT-2004 (TReMBLrel. 13, Last sequence update)  
 DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)  
 DE C63269-PA (LD07994P).  
 GN Name=Dmn; ORFNames=C63269;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN  
 RP [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Goebye J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Gallé R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazey R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA Abell J.P., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burdies K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Doonan K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz S., Ferriere S., Fleischmann W.,  
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,  
 RA Glodet A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibeagam C.,  
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kechum K.A.,  
 RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Lin X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Patel J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrtkars R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach U.,  
 RA Williams S.M., Woodgett M., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN  
 RP [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537572;  
 RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
 RA Patel J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svrtkars R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
 RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila  
 melanogaster euchromatic genome sequence.";  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN  
 RP [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kamlirer J.S., Bergman C.M., Krommiller B., Carlson J., Svrtkars R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the Drosophila melanogaster euchromatin:  
 a genomic perspective.";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).

RN  
 RP [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kamlirer J.S., Milburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield B.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Ruso S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN  
 RP [5]  
 RP SEQUENCE FROM N.A.  
 RX FLYbase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN  
 RP [6]  
 RP SEQUENCE FROM N.A.  
 RX FLYbase;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 RN  
 RP [7]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN-Berkeley;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Munoz J., Patel J., Paragas V., Park S., Phouanavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE003835; AAF59034.1; -  
 DR EMBL; AY061092; AAL28640.1; -  
 DR InterAct; 09V4Y9; -  
 DR FLYbase; FBgn0021825; Dmn.  
 DR GO; GO:0005869; C:dynactin complex; IEA.  
 DR GO; GO:0007017; P:microtubule-based process; IEA.  
 DR InterPro; IPR006996; Dynactin.  
 DR Pfam; PF04912; Dynactin; 1.  
 SQ SEQUENCE 380 AA; 41998 MW; CP7E1D3BFF5989C5 CRC64;  
 Query Match 28.9%; Score 72.5; DB 2; Length 380;  
 Best Local Similarity 48.7%; Pred. No. 34;  
 Matches 19; Conservative 5; Mismatches 14; Indels 1; Gaps 1;  
 C7 1 GVKETPOKXORLHVEQELTEVEKIKTYKESATEK 39  
 Db 94 GKEKTPVKCORLQIEMNELNEVALQVD-RKVADDEK 131  
 RESULT 11  
 ID Q6MVP7 PRELIMINARY; PRT; 800 AA.  
 AC Q6MVP7;  
 DT 05-JUL-2004 (TReMBLrel. 27, Created)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
 DE Probable translation elongation factor EF-G, mitochondrial.  
 GN Name=B16D18.090;  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 NCBI\_TaxID=5141;  
 RN  
 RP [1]  
 RP SEQUENCE FROM N.A.  
 RA Schulte U., Aign V., Hohnsels J., Brandt P., Fartmann B., Holland R.,  
 RA Nykatura G., Mewes H.W., Mannhaupt G.;  
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
 RN  
 RP [2]  
 RP SEQUENCE FROM N.A.  
 RA German Neurospora genome project;  
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BX842625; CAE76249.1; -  
 DR GO; GO:0005525; F:GTP binding; IEA.



```
DR GO: GO:0003746; F:translation elongation factor activity; IEA.
DR GO: GO:0006412; P:protein biosynthesis; IEA.
DR GO: GO:0006414; P:translational elongation; IEA.
DR InterPro: IPR004540; EF-G.
DR InterPro: IPR000640; ERG_C.
DR InterPro: IPR009022; ERG_III_V.
DR InterPro: IPR005517; ERG_IV_.
DR InterPro: IPR004161; EFTU_D2.
DR InterPro: IPR007951; ProTSyn_GTPbind.
DR InterPro: IPR005225; Small_GTP.
DR InterPro: IPR009000; Translat_factor.
DR Pfam: PF00679; ERG_C; 1.
DR Pfam: PF03764; ERG_IV; 1.
DR Pfam: PF00009; GTP_EFTU; 1.
DR Pfam: PF03144; GTP_EFTU_D2; 1.
DR PRINTS: PR00315; ELONGATINCT.
DR TIGRFAMs: TIGR00484; EF-G; 1.
DR TIGRFAMs: TIGR00231; small_GTP; 1.
DR PROSITE: PS00301; EFACOR_GTP; 1.
KW Elongation factor; GTP-binding; Protein biosynthesis.
SQ SEQUENCE 800 AA; 88373 MW; F2A7C425F6D4974A CRC64;

Query Match 28.3%; Score 71; DB 2; Length 800;
Best Local Similarity 26.2%; Pred. No. 1e+02;
Matches 16; Conservative 14; Mismatches 19; Indels 12; Gaps 1;

QY 2 VKETPOQKYORLLHEVOELTTEV-----EKIKTTVKESATEEKLTPVYLAKQL 49
DB 306 IKELCEQKQRELIKLVADVDDEIEMFLEBQTPPEQIKAIRATYACKTPTVLMGSAI 365

QY 50 A 50
DB 366 A 366

RESULT 12
Q7SH14 PRELIMINARY; PRT; 813 AA.
ID Q7SH14
AC Q7SH14;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=NCU02955.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=0874A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Imaniel P., Pedersen D., Nelson M., Washburne M.,
RA Selfetrenikoff C.P., Kinsey J.A., Braun E.L., Zeller A., Schulte U.,
RA Koche G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamyssejls M., Mauceli E., Bielke C., Rudd S., Friseman D.,
RA Krysiofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA Desouza C., Glase L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seltzer S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.",
RL Nature 0.0-0(2003).
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AABX0100006; EAA36106.1; -.
DR HSSP: P13551; IFNM.
DR GO: GO:0005525; F:GTP binding; IEA.
```

```
DR GO: GO:0003746; F:translation elongation factor activity; IEA.
DR GO: GO:0006412; P:protein biosynthesis; IEA.
DR GO: GO:0006414; P:translational elongation; IEA.
DR InterPro: IPR004540; EF-G.
DR InterPro: IPR000640; ERG_C.
DR InterPro: IPR009022; ERG_III_V.
DR InterPro: IPR005517; ERG_IV_.
DR InterPro: IPR004161; EFTU_D2.
DR InterPro: IPR007951; ProTSyn_GTPbind.
DR InterPro: IPR005225; Small_GTP.
DR InterPro: IPR009000; Translat_factor.
DR Pfam: PF00679; ERG_C; 1.
DR Pfam: PF03764; ERG_IV; 1.
DR Pfam: PF00009; GTP_EFTU; 1.
DR Pfam: PF03144; GTP_EFTU_D2; 1.
DR PRINTS: PR00315; ELONGATINCT.
DR TIGRFAMs: TIGR00484; EF-G; 1.
DR TIGRFAMs: TIGR00231; small_GTP; 1.
DR PROSITE: PS00301; EFACOR_GTP; 1.
KW GTP-binding; Hypothetical protein; Protein biosynthesis.
SQ SEQUENCE 813 AA; 89810 MW; C9A5093F35B29B05 CRC64;

Query Match 28.3%; Score 71; DB 2; Length 813;
Best Local Similarity 26.2%; Pred. No. 1e+02;
Matches 16; Conservative 14; Mismatches 19; Indels 12; Gaps 1;

QY 2 VKETPOQKYORLLHEVOELTTEV-----EKIKTTVKESATEEKLTPVYLAKQL 49
DB 306 IKELCEQKQRELIKLVADVDDEIEMFLEBQTPPEQIKAIRATYACKTPTVLMGSAI 365

QY 50 A 50
DB 366 A 366

RESULT 13
Q971V5 PRELIMINARY; PRT; 311 AA.
ID Q971V5
AC Q971V5;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein S71273.
GN OrderedLocustNames=S71273;
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX MEDLINE=21456156; PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Anrai A., Kosugi H., Hosoyma A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshikawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagita M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.",
RL DNA Res. 8:123-140(2001).
DR EMBL: AP000985; BAB66315.1; -.
DR HSSP: P01100; IFOS.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 311 AA; 37536 MW; 674338A5E031B429 CRC64;

Query Match 27.9%; Score 70; DB 2; Length 311;
Best Local Similarity 35.3%; Pred. No. 50;
Matches 18; Conservative 11; Mismatches 18; Indels 4; Gaps 2;

QY 2 VKETPOQKYORLLHEVOELTTEVEKIKTTVKESATEEKLTPVYLAKQLAAL 52
DB 71 IKOLSOQK-OSLIQVISEIKKEFEQIKNVEK---VKKLDPLQILKIKIEQL 117
```

```

RESULT 14
ID Q8MOK1 PRELIMINARY; PRT; 751 AA.
AC Q8MOK1;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE l008185p.
GN ORFNames=CG10971;
OS Drosophila melanogaster (Fruit Fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stadleton M., Brokstein P., Hong L., Abghayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Fafan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.U., Nuno J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phoumenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (JUL-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY129439; AAM76181.1; -.
DR FlyBase; FBgn0036309; CG10971.
DR GO; GO:0003779; F:actin binding; IEA.
DR InterPro; IPR002558; ILMQ.
DR Pfam; PF01608; I_LMEQ; 1.
DR PROSITE; PS50945; I_LMEQ; 1.
SQ SEQUENCE 751 AA; 85362 MW; 7EBC7661ECB8A0 CRC64;

Query Match 27.9%; Score 70; DB 2; Length 751;
Best Local Similarity 44.2%; Pred. NO. 1.2e+02;
Matches 19; Conservative 9; Mismatches 11; Indels 4; Gaps 2;

OY 7 QOKYORLLHEVOELTTEVEKIKTYKESATBEKLTPLVLAQOL 49
Db 72 KOVNSQLLTKELTKELTNEISKIKVNE--KPK-TNLTLOKQI 110

RESULT 15
ID Q8GBS5 PRELIMINARY; PRT; 1087 AA.
AC Q8GBS5;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE CG10971-PB.
GN ORFNames=CG10971;
OS Drosophila melanogaster (Fruit Fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN
RP SEQUENCE FROM N.A.
RL MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abail J.F., Abghayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Bendable J., Bayraktiroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Botkova D., Botchan M.R., Bouck Y., Brokstein P., Brottier P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

```

```

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foeller C., Gabriellian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
RA Jalali M., Kallush F., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milehina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stadleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter B., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers R.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
[2]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537558;
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Taber P.E., Wan K., Stadleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers R.W., Gibbs R.A., Rubin G.M.,
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).
[3]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomic perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).
[4]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktiroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
RA Stadleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083 (2002).
[5]
RN
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB003540; AAF49884.1; -.
DR FlyBase; FBgn0036309; CG10971.
DR GO; GO:0003779; F:actin binding; IEA.
DR GO; GO:0005543; F:phospholipid binding; IEA.
DR InterPro; IPR01417; ANTH.
DR InterPro; IPR001026; Epsin_N.

```



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: November 3, 2005, 21:51:39 ; Search time 26.6952 Seconds  
(without alignments)  
145,410 Million cell updates/sec

Title: 09782816-51

Perfect score: 251  
Sequence: 1 GVKETPOOKYQRLHEVQEL.....ESATFEKLTPLAKQIALAL 52

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/6D.COMB.pep:\*  
7: /cgn2\_6/ptodata/1/1aa/6E.COMB.pep:\*  
8: /cgn2\_6/ptodata/1/1aa/6F.COMB.pep:\*  
9: /cgn2\_6/ptodata/1/1aa/6G.COMB.pep:\*  
10: /cgn2\_6/ptodata/1/1aa/6H.COMB.pep:\*  
11: /cgn2\_6/ptodata/1/1aa/6I.COMB.pep:\*  
12: /cgn2\_6/ptodata/1/1aa/6J.COMB.pep:\*  
13: /cgn2\_6/ptodata/1/1aa/6K.COMB.pep:\*  
14: /cgn2\_6/ptodata/1/1aa/6L.COMB.pep:\*  
15: /cgn2\_6/ptodata/1/1aa/6M.COMB.pep:\*  
16: /cgn2\_6/ptodata/1/1aa/6N.COMB.pep:\*  
17: /cgn2\_6/ptodata/1/1aa/6O.COMB.pep:\*  
18: /cgn2\_6/ptodata/1/1aa/6P.COMB.pep:\*  
19: /cgn2\_6/ptodata/1/1aa/6Q.COMB.pep:\*  
20: /cgn2\_6/ptodata/1/1aa/6R.COMB.pep:\*  
21: /cgn2\_6/ptodata/1/1aa/6S.COMB.pep:\*  
22: /cgn2\_6/ptodata/1/1aa/6T.COMB.pep:\*  
23: /cgn2\_6/ptodata/1/1aa/6U.COMB.pep:\*  
24: /cgn2\_6/ptodata/1/1aa/6V.COMB.pep:\*  
25: /cgn2\_6/ptodata/1/1aa/6W.COMB.pep:\*  
26: /cgn2\_6/ptodata/1/1aa/6X.COMB.pep:\*  
27: /cgn2\_6/ptodata/1/1aa/6Y.COMB.pep:\*  
28: /cgn2\_6/ptodata/1/1aa/6Z.COMB.pep:\*  
29: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
30: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
31: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep:\*  
32: /cgn2\_6/ptodata/1/1aa/6D.COMB.pep:\*  
33: /cgn2\_6/ptodata/1/1aa/6E.COMB.pep:\*  
34: /cgn2\_6/ptodata/1/1aa/6F.COMB.pep:\*  
35: /cgn2\_6/ptodata/1/1aa/6G.COMB.pep:\*  
36: /cgn2\_6/ptodata/1/1aa/6H.COMB.pep:\*  
37: /cgn2\_6/ptodata/1/1aa/6I.COMB.pep:\*  
38: /cgn2\_6/ptodata/1/1aa/6J.COMB.pep:\*  
39: /cgn2\_6/ptodata/1/1aa/6K.COMB.pep:\*  
40: /cgn2\_6/ptodata/1/1aa/6L.COMB.pep:\*  
41: /cgn2\_6/ptodata/1/1aa/6M.COMB.pep:\*  
42: /cgn2\_6/ptodata/1/1aa/6N.COMB.pep:\*  
43: /cgn2\_6/ptodata/1/1aa/6O.COMB.pep:\*  
44: /cgn2\_6/ptodata/1/1aa/6P.COMB.pep:\*  
45: /cgn2\_6/ptodata/1/1aa/6Q.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66	26.3	142	4	US-09-270-767-43370
2	64.5	25.7	639	4	US-09-902-540-14908
3	64	25.5	131	4	US-09-513-999C-7806
4	64	25.5	193	4	US-09-107-532A-5587
5	63	25.1	850	4	US-09-269-858A-8
6	63	25.1	851	4	US-09-623-326-43
7	62.5	24.9	1270	4	US-09-538-092-1321
8	61.5	24.5	309	4	US-09-248-786A-14974
9	61	24.3	442	3	US-09-081-689-2
10	61	24.3	442	3	US-09-305-984-14
11	61	24.3	442	3	US-09-073-541A-14
12	61	24.3	442	4	US-09-493-940-14
13	61	24.3	442	4	US-09-583-110-5317
14	61	24.3	446	4	US-09-107-433-3720
15	61	24.3	644	1	US-08-487-890A-6
16	61	24.3	644	1	US-08-478-435-6
17	61	24.3	644	2	US-08-337-483-6
18	61	24.3	644	2	US-08-478-373-6
19	61	24.3	644	3	US-08-474-671-6
20	61	24.3	644	3	US-08-483-577A-6
21	61	24.3	644	3	US-08-897-438-6
22	61	24.3	644	3	US-08-637-654-6
23	61	24.3	644	3	US-08-649-518-6
24	60.5	24.1	87	3	US-08-284-033-3
25	60.5	24.1	87	3	US-08-729-834B-3
26	60.5	24.1	1217	4	US-09-949-016-7454
27	59.5	23.7	424	4	US-09-286-981B-14

28	59.5	23.7	425	4	US-09-286-981B-13	Sequence 13, Appl
29	59.5	23.7	426	4	US-09-286-981B-12	Sequence 12, Appl
30	59.5	23.7	721	4	US-09-248-786A-18800	Sequence 18800, A
31	59	23.5	451	4	US-09-270-767-45139	Sequence 45139, A
32	59	23.5	2101	1	US-08-466-390-4	Sequence 4, Appl
33	59	23.5	2101	1	US-08-470-950-4	Sequence 4, Appl
34	59	23.5	2101	1	US-08-467-781-4	Sequence 4, Appl
35	59	23.5	2101	1	US-08-195-487-4	Sequence 4, Appl
36	59	23.5	2101	2	US-08-483-924-4	Sequence 4, Appl
37	59	23.5	2101	3	US-09-452-294-1	Sequence 1, Appl
38	59	23.5	2101	3	US-09-452-294-1	Sequence 1, Appl
39	59	23.5	2107	4	US-09-949-016-7646	Sequence 7646, Ap
40	59	23.5	2107	4	US-09-949-016-7647	Sequence 7647, Ap
41	58.5	23.3	399	3	US-09-284-033-2	Sequence 2, Appl
42	58.5	23.3	399	3	US-08-729-834B-2	Sequence 2, Appl
43	58.5	23.3	399	3	US-09-949-016-6466	Sequence 6466, Ap
44	58.5	23.3	404	4	US-09-949-016-8325	Sequence 8325, Ap
45	58	23.1	487	3	US-09-724-224-8	Sequence 8, Appl

## ALIGNMENTS

RESULT 1  
US-09-270-767-43370  
Sequence 43370, Application US/09270767  
Patent No. 6703491  
GENERAL INFORMATION:  
APPLICANT: Homburger et al.  
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
FILE REFERENCE: File Reference: 7326-094  
CURRENT FILING DATE: 1999-03-17  
NUMBER OF SEQ ID NOS: 62517  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 43370  
LENGTH: 142  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-09-270-767-43370

Query Match 26.3%; Score 66; DB 4; Length 142;  
Best Local Similarity 50.0%; Pred. No. 1.6;  
Matches 16; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

1 GVKETPOOKYQRLHEVQELTEVEKIKTKYK 32  
Db 111 GVKETPOOKYQRLHEVQELTEVEKIKTKYK 142

RESULT 2  
US-09-902-540-14908  
Sequence 14908, Application US/09902540  
Patent No. 6833447  
GENERAL INFORMATION:  
APPLICANT: Goldman, Barry S.  
TITLE OF INVENTION: Hinkle, Gregory J.  
FILE REFERENCE: 38-10(11549)B  
CURRENT FILING DATE: 2001-07-10  
CURRENT FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: 60/217,883  
PRIOR FILING DATE: 2000-07-10  
NUMBER OF SEQ ID NOS: 16825  
SEQ ID NO 14908  
LENGTH: 639  
TYPE: PRT  
ORGANISM: Myxococcus xanthus  
US-09-902-540-14908

Query Match 25.7%; Score 64.5; DB 4; Length 639;

Best Local Similarity 31.9%; Pred. No. 14;  
Matches 15; Conservative 11; Mismatches 12; Indels 9; Gaps 1;

Qy 1 GVKETPOQKQRLHVEQELTTEVERK-----IKTTVKESATEE 38  
Db 151 GALEAVOEQYOQAIADVERLTSELEAASAEKDSIGLRTQLEALAE 197

## RESULT 3

US-09-513-999C-7806  
Sequence 7806, Application US/09513999C  
Patent No. 6783961  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Duglert, A.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
Patent No. 6783961  
FILE REFERENCE: 59, US2, REG  
CURRENT APPLICATION NUMBER: US/09/513,999C  
CURRENT FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/122,487  
PRIOR FILING DATE: 1999-02-26  
NUMBER OF SEQ ID NOS: 36681  
SOFTWARE: Patent.pm  
SEQ ID NO 7806  
LENGTH: 151  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SIGNAL  
LOCATION: -25...-1  
OTHER INFORMATION: score 3.9  
OTHER INFORMATION: seq RYGLLLGGGGVYG/SR  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: 94  
OTHER INFORMATION: Xaa=Ile or Lys or Asn or Arg or Ser or Thr  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: 103  
OTHER INFORMATION: Xaa=Glu or Lys  
US-09-513-999C-7806

Query Match 25.5%; Score 64; DB 4; Length 151;  
Best Local Similarity 40.5%; Pred. No. 3;

Matches 17; Conservative 7; Mismatches 8; Indels 10; Gaps 1;

Qy 9 KYORLHVEQELTTEVERKIKTTVKESA-----TEEKL 40  
Db 86 KYRKLKAEVKKSKLEKKKKTITTSARQKQKXIERQEK 127

## RESULT 4

US-09-107-532A-5587  
Sequence 5587, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:  
NAME: Arindello, Pamela Deneke

REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 5587:

SEQUENCE CHARACTERISTICS:  
LENGTH: 193 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium

FEATURE:  
NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...193

SEQUENCE DESCRIPTION: SEQ ID NO: 5587:

US-09-107-532A-5587

Query Match 25.5%; Score 64; DB 4; Length 193;  
Best Local Similarity 40.5%; Pred. No. 3.9;

Matches 15; Conservative 6; Mismatches 14; Indels 2; Gaps 1;

Qy 2 VKETPOQKQRLHVEQELTTEVERKIKTTVKESATEE 38  
Db 70 VSEFVEQTVBELER--EPTAEIEVNVETILEEKTEE 104

## RESULT 5

US-09-269-858A-8  
Sequence 8, Application US/09269858A  
Patent No. 6692932  
GENERAL INFORMATION:  
APPLICANT: Angerer, Bernhard  
APPLICANT: Ankenbauer, Waltraud  
APPLICANT: Boehringer, M.  
APPLICANT: Bonch-Osmolovskaya, Elizaveta  
APPLICANT: Markau, Ursula  
APPLICANT: Reiser, Astrid  
APPLICANT: Schmitz-Agheguyan, Gudrun  
TITLE OF INVENTION: Thermostable DNA Polymerase from Anaerocellum Thermophilum  
FILE REFERENCE: 4452  
CURRENT APPLICATION NUMBER: US/09/269,858A  
PRIOR APPLICATION NUMBER: PCT/EP97/05390  
PRIOR FILING DATE: 1996-10-03  
PRIOR APPLICATION NUMBER: EP/96115877  
PRIOR FILING DATE: 1996-10-03  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 8  
LENGTH: 850  
TYPE: PRT  
ORGANISM: Abedus herberti  
US-09-269-858A-8

Query Match 25.1%; Score 63; DB 4; Length 850;  
Best Local Similarity 30.9%; Pred. No. 28;

Matches 17; Conservative 15; Mismatches 19; Indels 4; Gaps 2;

Qy 1 GVKETPOOKVORLHE---VOELTVEVKITVYKESATEERKLPVLAKROL 52  
Db 192 GVKIGIKSKAKLLEYSLSLEIYQNDIKIKSIREKLEAGK-DMAFLSKRLATI 245

RESULT 6  
US-09-623-326-43  
; Sequence 43, Application US/09623326  
; Patent No. 6607883  
; GENERAL INFORMATION:  
; APPLICANT: Frey et al.  
; TITLE OF INVENTION: Polymerase Chimerae  
; FILE REFERENCE: 4894  
; CURRENT APPLICATION NUMBER: US/09/623,326  
; CURRENT FILING DATE: 2001-02-08  
; PRIOR APPLICATION NUMBER: DE 198 10 879.6  
; PRIOR FILING DATE: 1998-03-13  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 43  
; LENGTH: 851  
; TYPE: PRT  
; ORGANISM: Anaerocellum thermophilum  
US-09-623-326-43

Query Match 25.1%; Score 63; DB 4; Length 851;  
Best Local Similarity 30.9%; Pred. No. 28;  
Matches 17; Conservative 15; Mismatches 19; Indels 4; Gaps 2;

Qy 1 GVKETPOOKVORLHE---VOELTVEVKITVYKESATEERKLPVLAKROL 52  
Db 192 GVKIGIKSKAKLLEYSLSLEIYQNDIKIKSIREKLEAGK-DMAFLSKRLATI 245

RESULT 7  
US-09-538-092-1321  
; Sequence 1321, Application US/09538092  
; Patent No. 6753314  
; GENERAL INFORMATION:  
; APPLICANT: Glot, Loic  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538,092  
; CURRENT FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/178,965  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CuratseqFormatter Version 0.9  
; SEQ ID NO 1321  
; LENGTH: 1270  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (0)...(0)  
; OTHER INFORMATION: Polypeptide Accession Number Q14203  
US-09-538-092-1321

Query Match 24.9%; Score 62.5; DB 4; Length 1270;  
Best Local Similarity 34.8%; Pred. No. 52;  
Matches 16; Conservative 11; Mismatches 18; Indels 1; Gaps 1;

Qy 4 ETPOOKVORLHEVOELTVEVKITVYKESATEERKLPVLAKROL 49  
Db 311 ESLQOEVBALKERYDELTTDLIIKAEIEKSGDAASVQL-KOL 355

RESULT 8  
US-09-248-796A-14974  
; Sequence 14974, Application US/09248796A

; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Kelch Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN;  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 14974  
; LENGTH: 309  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-14974

Query Match 24.5%; Score 61.5; DB 4; Length 309;  
Best Local Similarity 41.2%; Pred. No. 13;  
Matches 14; Conservative 8; Mismatches 9; Indels 3; Gaps 1;

Qy 3 KETPO---KQORLHEVOELTVEVKITVKE 33  
Db 271 EETPIQDNPKRRIHBIIDTSDIEDIETPTPE 304

RESULT 9  
US-09-081-689-2  
; Sequence 2, Application US/09081689  
; Patent No. 6165992  
; GENERAL INFORMATION:  
; APPLICANT: Wallis, Nicola G.  
; APPLICANT: Zalacain, Magdalena  
; APPLICANT: Throup, John  
; APPLICANT: Biswas, Sanjoy  
; TITLE OF INVENTION: Histidine Kinase  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechart, Price & Rhoads  
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103-2793  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/081,689  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/048,347  
; FILING DATE: 30-MAY-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Falk, Stephen T  
; REGISTRATION NUMBER: 36,795  
; REFERENCE/DOCKET NUMBER: GML0009  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-994-2488  
; TELEFAX: 215-994-2222  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 442 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-081-689-2

```
Query Match      24.3% Score 61; DB 3; Length 442;
Best Local Similarity 33.3% Pred. No. 23;
Matches 18; Conservative 6; Mismatches 24; Indels 6; Gaps 1;

Qy      2 VKETPQOKYQRL-----LHEVOELTTEVEKIKTTVESATEEKLTPVLLAKOL 49
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db      200 LKEQINSLYQHLLTVIADLHEKNEAIIQLQEKMKVFLRGASHELKTPLASLKITL 253

RESULT 10
US-09-984-14
; Sequence 14, Application US/09305984B
; Patent No. 6331407
; GENERAL INFORMATION:
; APPLICANT: No. 6331407ak, Rodger
; TITLE OF INVENTION: NOVEL ANTIBIOTICS AND METHODS OF USING THE SAME
; FILE REFERENCE: 1340-1-016N1
; CURRENT APPLICATION NUMBER: US/09/305,984B
; EARLIER FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: 60/084,399
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: 09/305,984
; EARLIER FILING DATE: 1999-05-05
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 14
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-305-984-14

Query Match      24.3% Score 61; DB 3; Length 442;
Best Local Similarity 33.3% Pred. No. 23;
Matches 18; Conservative 6; Mismatches 24; Indels 6; Gaps 1;

Qy      2 VKETPQOKYQRL-----LHEVOELTTEVEKIKTTVESATEEKLTPVLLAKOL 49
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db      200 LKEQINSLYQHLLTVIADLHEKNEAIIQLQEKMKVFLRGASHELKTPLASLKITL 253

RESULT 11
US-09-073-541A-14
; Sequence 14, Application US/09073541A
; Patent No. 6448224
; GENERAL INFORMATION:
; APPLICANT: No. 6448224ak, Rodger
; TITLE OF INVENTION: NOVEL ANTIBIOTICS AND METHODS OF USING THE SAME
; FILE REFERENCE: 1340-1-016
; CURRENT APPLICATION NUMBER: US/09/073,541A
; EARLIER FILING DATE: 1998-05-06
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 14
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-073-541A-14

Query Match      24.3% Score 61; DB 4; Length 442;
Best Local Similarity 33.3% Pred. No. 23;
Matches 18; Conservative 6; Mismatches 24; Indels 6; Gaps 1;

Qy      2 VKETPQOKYQRL-----LHEVOELTTEVEKIKTTVESATEEKLTPVLLAKOL 49
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db      200 LKEQINSLYQHLLTVIADLHEKNEAIIQLQEKMKVFLRGASHELKTPLASLKITL 253

RESULT 12
US-09-493-940-14
; Sequence 14, Application US/09493940

Patent No. 6630583
; GENERAL INFORMATION:
; APPLICANT: No. 6630583ak, Rodger
; APPLICANT: Toumanen, Elaine
; TITLE OF INVENTION: NOVEL ANTIBIOTICS AND METHODS OF USING THE SAME
; FILE REFERENCE: 1340-1-016N1
; CURRENT APPLICATION NUMBER: US/09/493,940
; EARLIER FILING DATE: 2000-01-28
; EARLIER APPLICATION NUMBER: 60/084,399
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: 09/305,984
; EARLIER FILING DATE: 1999-05-05
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 14
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-493-940-14

Query Match      24.3% Score 61; DB 4; Length 442;
Best Local Similarity 33.3% Pred. No. 23;
Matches 18; Conservative 6; Mismatches 24; Indels 6; Gaps 1;

Qy      2 VKETPQOKYQRL-----LHEVOELTTEVEKIKTTVESATEEKLTPVLLAKOL 49
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db      200 LKEQINSLYQHLLTVIADLHEKNEAIIQLQEKMKVFLRGASHELKTPLASLKITL 253

RESULT 13
US-09-583-110-5317
; Sequence 5317, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PAT00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; EARLIER FILING DATE: 2000-05-26
; EARLIER APPLICATION NUMBER: US 09/107,433
; EARLIER FILING DATE: 1998-06-30
; EARLIER APPLICATION NUMBER: US 60/085,131
; EARLIER FILING DATE: 1998-05-12
; EARLIER APPLICATION NUMBER: US 60/051,553
; EARLIER FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 5317
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-5317

Query Match      24.3% Score 61; DB 4; Length 442;
Best Local Similarity 33.3% Pred. No. 23;
Matches 18; Conservative 6; Mismatches 24; Indels 6; Gaps 1;

Qy      2 VKETPQOKYQRL-----LHEVOELTTEVEKIKTTVESATEEKLTPVLLAKOL 49
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db      200 LKEQINSLYQHLLTVIADLHEKNEAIIQLQEKMKVFLRGASHELKTPLASLKITL 253

RESULT 14
US-09-107-433-3720
; Sequence 3720, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
```



ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Maitiam  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: <Unknown>  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: <Unknown>  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,433  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085131  
FILING DATE: May 12, 1998  
APPLICATION NUMBER: 60/051553  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Arinello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-011  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 3720:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 446 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pneumoniae  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (8) LOCATION 1...446  
SEQUENCE DESCRIPTION: SEQ ID NO: 3720:  
US-09-107-433-3720

Query March 24.3%; Score 61; DB 4; Length 446;  
Best Local Similarity 33.3%; Pred. No. 23;  
Matches 18; Conservative 6; Mismatches 24; Indels 6; Gaps 1;

Qy 2 VKETPOQKYQRL-----LHEVOELTTEVKIKTTVKESATEBKLTPLVLAQOL 49  
Db 204 LKEQINSIYQHLVVIADLHKRAIILQLEKMYEFLRGASHKLTPLASIKIL 257

RESULT 15  
US-08-487-890A-6  
Sequence 6, Application US/08487890A  
Patent No. 5708149  
GENERAL INFORMATION:  
APPLICANT: Loosmore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Mirdin, Andrew  
APPLICANT: Klein, Michael  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,890A  
FILING DATE: 07-JUN-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/175,116  
FILING DATE: 29-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,968  
FILING DATE: 08-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-466 MTS:jd  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 644 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-487-890A-6

Query March 24.3%; Score 61; DB 1; Length 644;  
Best Local Similarity 32.6%; Pred. No. 36;  
Matches 15; Conservative 8; Mismatches 17; Indels 6; Gaps 1;

Qy 2 VKETPOQKYQRLHEVOELTTEVE-----KIKTTVKESATEBKLT 41  
Db 244 LRETDANKSQNRTHKLVDLEADVSNRFRGKVRKTKESSEHPT 289

Search completed: November 3, 2005, 22:05:57  
Job time: 27.6952 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: November 3, 2005, 21:53:15 ; Search time 93.7112 Seconds  
(without alignments)  
232.174 Million cell updates/sec

Title: 09782816-51

Perfect score: 251

Sequence: 1 GVKEETPOOKYRQLLHEVDEL.....ESATEKLTPLVLLAKQLAAL 52

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US05\_PUBCOMB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US04\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US03\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US02\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US01\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US00\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US00A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US00B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US00C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US00D\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US00E\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US00F\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US00G\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US00H\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US00I\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US00J\_PUBCOMB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US00K\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US00L\_PUBCOMB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US00M\_PUBCOMB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US00N\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	251	100.0	401	16	US-10-408-765A-1369
2	251	100.0	406	16	US-09-782-816A-53
3	251	100.0	465	9	US-09-925-298-676
4	251	100.0	465	14	US-10-102-806-676
5	248	98.8	183	10	US-09-782-816A-54
6	248	98.8	224	16	US-10-425-115-315831
7	246	98.0	52	10	US-09-782-816A-51
8	231	92.0	134	14	US-10-106-698-6730
9	113	45.0	22	10	US-09-782-816A-3
10	107	42.6	21	10	US-09-782-816A-4
11	103	41.0	20	10	US-09-782-816A-5

	12	103	41.0	23	10	US-09-782-816A-1	Sequence 1, Appli
	13	98	35.0	19	10	US-09-782-816A-6	Sequence 6, Appli
	14	93	37.1	18	10	US-09-782-816A-7	Sequence 7, Appli
	15	88	35.1	17	10	US-09-782-816A-8	Sequence 8, Appli
	16	81	32.3	16	10	US-09-782-816A-9	Sequence 9, Appli
	17	76	30.3	15	10	US-09-782-816A-10	Sequence 10, Appli
	18	72.5	28.9	53	10	US-09-782-816A-52	Sequence 52, Appli
	19	72.5	28.9	380	10	US-09-782-816A-56	Sequence 56, Appli
	20	72.5	28.9	380	20	US-11-097-143-4056	Sequence 4056, Ap
	21	71	28.3	14	10	US-09-782-816A-11	Sequence 11, Appli
	22	70.5	28.1	348	16	US-10-425-115-110465	Sequence 210465,
	23	70.5	28.1	356	15	US-10-425-114-63356	Sequence 63356, A
	24	70	27.9	1087	20	US-11-097-143-22866	Sequence 22866, A
	25	70	27.9	1087	20	US-11-097-143-22866	Sequence 22866, A
	26	68	27.1	2368	9	US-09-815-242-5635	Sequence 5635, Ap
	27	68	27.1	2368	9	US-09-815-242-12389	Sequence 12389, A
	28	67	26.7	2478	9	US-09-815-242-5816	Sequence 5816, Ap
	29	67	26.7	2478	9	US-09-815-242-12967	Sequence 12967, A
	30	67	26.7	2478	17	US-10-470-048B-220	Sequence 220, App
	31	67	26.7	2481	15	US-10-282-122A-43762	Sequence 43762, A
	32	66	26.3	13	10	US-09-782-816A-12	Sequence 12, Appli
	33	66	26.3	386	15	US-10-282-122A-71967	Sequence 71967, A
	34	65.5	26.1	576	15	US-10-424-599-274041	Sequence 274041, A
	35	65.5	26.1	1837	15	US-10-369-493-22734	Sequence 22734, A
	36	64.5	25.7	348	16	US-10-437-963-139713	Sequence 139713, A
	37	64	25.5	185	16	US-10-842-740-3	Sequence 3, Appli
	38	64	25.5	188	10	US-09-284-320-12	Sequence 12, Appli
	39	64	25.5	188	14	US-10-028-072-2	Sequence 2, Appli
	40	64	25.5	188	14	US-10-140-808-2	Sequence 2, Appli
	41	64	25.5	188	14	US-10-121-049-2	Sequence 2, Appli
	42	64	25.5	188	14	US-10-123-904-2	Sequence 2, Appli
	43	64	25.5	188	14	US-10-140-470-2	Sequence 2, Appli
	44	64	25.5	188	14	US-10-175-746-2	Sequence 2, Appli
	45	64	25.5	188	14	US-10-176-918-2	Sequence 2, Appli

#### ALIGNMENTS

RESULT 1  
US-10-408-765A-1369  
; Sequence 1369, Application US/10408765A  
; Publication No. US20040101874A1  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Fahy, Bojn D.  
; APPLICANT: Zhang, Bing  
; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.  
; APPLICANT: Glenn, Gary W.  
; APPLICANT: Warnock, Dale E.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
; FILE REFERENCE: 660088.465  
; CURRENT APPLICATION NUMBER: US/10/408,765A  
; CURRENT FILING DATE: 2003-04-04  
; NUMBER OF SEQ ID NOS: 3077  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1369  
; LENGTH: 401  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-408-765A-1369

Query Match 100.0%; Score 251; DB 16; Length 401;  
Best Local Similarity 100.0%; Pred. No. 1; Se-19;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVKEETPOOKYRQLLHEVDELTFVEKIKTVKESATEKLTPLVLLAKQLAAL 52  
Db 94 GVKEETPOOKYRQLLHEVDELTFVEKIKTVKESATEKLTPLVLLAKQLAAL 145

RESULT 2  
US-09-782-816A-53  
; Sequence 53, Application US/09782816A  
; Publication No. US20030032771A1  
; GENERAL INFORMATION:  
; APPLICANT: Sharp, David J.  
; APPLICANT: Rogers, Gregory C.  
; APPLICANT: Scholey, Jonathan M.  
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR  
; TITLE OF INVENTION: PROLIFERATION  
; FILE REFERENCE: U0069.001A  
; CURRENT APPLICATION NUMBER: US/09/782.816A  
; CURRENT FILING DATE: 2001-02-14  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 53  
; LENGTH: 406  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-782-816A-53

Query Match 100.0%; Score 251; DB 10; Length 406;  
Best Local Similarity 100.0%; Pred. No. 1.6e-19;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GVKETPOOKYORLHVEQELTTEVEKIKITVKSATEEKLTPVLLAKQALAL 52  
DB 99 GVKETPOOKYORLHVEQELTTEVEKIKITVKSATEEKLTPVLLAKQALAL 150

RESULT 3  
US-09-925-298-676  
; Sequence 676, Application US/09925298  
; Publication No. US20020039764A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA103  
; CURRENT APPLICATION NUMBER: US/09/925.298  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05881  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 846  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 676  
; LENGTH: 465  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (5)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (6)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (16)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-298-676

Query Match 100.0%; Score 251; DB 9; Length 465;  
Best Local Similarity 100.0%; Pred. No. 1.8e-19;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GVKETPOOKYORLHVEQELTTEVEKIKITVKSATEEKLTPVLLAKQALAL 52  
DB 158 GVKETPOOKYORLHVEQELTTEVEKIKITVKSATEEKLTPVLLAKQALAL 209

RESULT 4  
US-10-102-806-676

; Sequence 676, Application US/10102806  
; Publication No. US20030054421A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA103P1C1  
; CURRENT APPLICATION NUMBER: US/10/102.806  
; CURRENT FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: 09/925,298  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05881  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 846  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 676  
; LENGTH: 465  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (5)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (6)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (16)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-102-806-676

Query Match 100.0%; Score 251; DB 14; Length 465;  
Best Local Similarity 100.0%; Pred. No. 1.8e-19;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GVKETPOOKYORLHVEQELTTEVEKIKITVKSATEEKLTPVLLAKQALAL 52  
DB 158 GVKETPOOKYORLHVEQELTTEVEKIKITVKSATEEKLTPVLLAKQALAL 209

RESULT 5  
US-09-782-816A-54  
; Sequence 54, Application US/09782816A  
; Publication No. US20030032771A1  
; GENERAL INFORMATION:  
; APPLICANT: Sharp, David J.  
; APPLICANT: Rogers, Gregory C.  
; APPLICANT: Scholey, Jonathan M.  
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR  
; TITLE OF INVENTION: PROLIFERATION  
; FILE REFERENCE: U0069.001A  
; CURRENT APPLICATION NUMBER: US/09/782.816A  
; CURRENT FILING DATE: 2001-02-14  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 54  
; LENGTH: 183  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-782-816A-54

Query Match 98.8%; Score 248; DB 10; Length 183;  
Best Local Similarity 98.1%; Pred. No. 1.3e-19;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GVKETPOOKYORLHVEQELTTEVEKIKITVKSATEEKLTPVLLAKQALAL 52  
DB 94 GVKETPOOKYORLHVEQELTTEVEKIKITVKSATEEKLTPVLLAKQALAL 145

RESULT 6  
US-10-425-115-315831

```
; Sequence 315831, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yinhua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 315831
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT457_51109C.1.pep
US-10-425-115-315831

Query Match          98.8%; Score 248; DB 16; Length 224;
Best Local Similarity 98.1%; Pred. No. 1.7e-19;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVKETPOOKYORLHVEQELTTEVEKIKTYVESATEBEKLTPIVLLAKQLAAL 52
DB 47 GVKETPOOKYORLHVEQELTTEVEKIKTYVESATEBEKLTPIVLLAKQLAAL 98

RESULT 7
US-09-782-816A-51
; Sequence 51, Application US/09782816A
; Publication No. US20030032771A1
; GENERAL INFORMATION:
; APPLICANT: Sharp, David J.
; APPLICANT: Rogers, Gregory C.
; APPLICANT: Scholey, Jonathan M.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
; TITLE OF INVENTION: PROLIFERATION
; FILE REFERENCE: UC069,001A
; CURRENT APPLICATION NUMBER: US/09/782,816A
; CURRENT FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 44
; OTHER INFORMATION: Xaa = Val or Leu
; OTHER INFORMATION: A sequence conserved among Homo sapiens and Mus
; OTHER INFORMATION: musculus.
US-09-782-816A-51

Query Match          98.0%; Score 246; DB 10; Length 52;
Best Local Similarity 98.1%; Pred. No. 5e-20;
Matches 51; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6730
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (126)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-6730

Query Match          92.0%; Score 231; DB 14; Length 134;
Best Local Similarity 100.0%; Pred. No. 7.1e-18;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TPQOKYORLHVEQELTTEVEKIKTYVESATEBEKLTPIVLLAKQLAAL 52
DB 1 TPQOKYORLHVEQELTTEVEKIKTYVESATEBEKLTPIVLLAKQLAAL 48

RESULT 9
US-09-782-816A-3
; Sequence 3, Application US/09782816A
; Publication No. US20030032771A1
; GENERAL INFORMATION:
; APPLICANT: Sharp, David J.
; APPLICANT: Rogers, Gregory C.
; APPLICANT: Scholey, Jonathan M.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
; TITLE OF INVENTION: PROLIFERATION
; FILE REFERENCE: UC069,001A
; CURRENT APPLICATION NUMBER: US/09/782,816A
; CURRENT FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: A sequence conserved among Homo sapiens and Mus
; OTHER INFORMATION: musculus.
US-09-782-816A-3

Query Match          45.0%; Score 113; DB 10; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVKETPOOKYORLHVEQELTT 22
DB 1 GVKETPOOKYORLHVEQELTT 22

RESULT 10
US-09-782-816A-4
; Sequence 4, Application US/09782816A
; Publication No. US20030032771A1
; GENERAL INFORMATION:
; APPLICANT: Sharp, David J.
; APPLICANT: Rogers, Gregory C.
; APPLICANT: Scholey, Jonathan M.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
; TITLE OF INVENTION: PROLIFERATION
```

```
FILE REFERENCE: UC069.001A
CURRENT APPLICATION NUMBER: US/09/782,816A
CURRENT FILING DATE: 2001-02-14
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 21
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: A sequence conserved among Homo sapiens and Mus
OTHER INFORMATION: musculus.
US-09-782-816A-4
```

```
Query Match 42.6%; Score 107; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.7e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2 VKETPOOKYQRLHVEQELTT 22
Db 1 VKETPOOKYQRLHVEQELTT 21
```

```
RESULT 11
US-09-782-816A-5
Sequence 5, Application US/09782816A
Publication No. US20030032771A1
GENERAL INFORMATION:
APPLICANT: Sharp, David J.
APPLICANT: Rogers, Gregory C.
APPLICANT: Scholey, Jonathan M.
TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
FILE REFERENCE: UC069.001A
CURRENT APPLICATION NUMBER: US/09/782,816A
CURRENT FILING DATE: 2001-02-14
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 20
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: A sequence conserved among Homo sapiens and Mus
OTHER INFORMATION: musculus.
US-09-782-816A-5
```

```
Query Match 41.0%; Score 103; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 3 KETPOOKYQRLHVEQELTT 22
Db 1 KETPOOKYQRLHVEQELTT 20
```

```
RESULT 12
US-09-782-816A-1
Sequence 1, Application US/09782816A
Publication No. US20030032771A1
GENERAL INFORMATION:
APPLICANT: Sharp, David J.
APPLICANT: Rogers, Gregory C.
APPLICANT: Scholey, Jonathan M.
TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
FILE REFERENCE: UC069.001A
CURRENT APPLICATION NUMBER: US/09/782,816A
CURRENT FILING DATE: 2001-02-14
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 23
```

```
TYPE: PRT
ORGANISM: Unknown
FEATURE:
NAME/KEY: UNSURE
LOCATION: 22
OTHER INFORMATION: Xaa = Val or Leu
OTHER INFORMATION: The sequence is a Homo sapiens sequence when Xaa
OTHER INFORMATION: represents Leu and a Mus musculus sequence when
OTHER INFORMATION: Xaa represents Val.
US-09-782-816A-1
```

```
Query Match 41.0%; Score 103; DB 10; Length 23;
Best Local Similarity 95.7%; Pred. No. 0.00015;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 23 EVEKIKTVKESATEEKLTPVLL 45
Db 1 EVEKIKTVKESATEEKLTPVXL 23
```

```
RESULT 13
US-09-782-816A-6
Sequence 6, Application US/09782816A
Publication No. US20030032771A1
GENERAL INFORMATION:
APPLICANT: Sharp, David J.
APPLICANT: Rogers, Gregory C.
APPLICANT: Scholey, Jonathan M.
TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
FILE REFERENCE: UC069.001A
CURRENT APPLICATION NUMBER: US/09/782,816A
CURRENT FILING DATE: 2001-02-14
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 19
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: A sequence conserved among Homo sapiens and Mus
OTHER INFORMATION: musculus.
US-09-782-816A-6
```

```
Query Match 39.0%; Score 98; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 4 ETPPOOKYQRLHVEQELTT 22
Db 1 ETPPOOKYQRLHVEQELTT 19
```

```
RESULT 14
US-09-782-816A-7
Sequence 7, Application US/09782816A
Publication No. US20030032771A1
GENERAL INFORMATION:
APPLICANT: Sharp, David J.
APPLICANT: Rogers, Gregory C.
APPLICANT: Scholey, Jonathan M.
TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
FILE REFERENCE: UC069.001A
CURRENT APPLICATION NUMBER: US/09/782,816A
CURRENT FILING DATE: 2001-02-14
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 18
TYPE: PRT
ORGANISM: Unknown
FEATURE:
```

OTHER INFORMATION: A sequence conserved among Homo sapiens and Mus  
OTHER INFORMATION: musculus.  
US-09-782-816A-7

Query Match 37.1%; Score 93; DB 10; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TPQKQRLHVEVQLTT 22  
|||||  
Db 1 TPQKQRLHVEVQLTT 18

RESULT 15  
US-09-782-816A-8  
Sequence 8, Application US/09782816A  
Publication No. US20030032771A1  
GENERAL INFORMATION:

APPLICANT: Sharp, David J.  
APPLICANT: Scholey, Gregory C.  
APPLICANT: Scholey, Jonathan M.  
TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR  
FILE REFERENCE: UC069,001A  
CURRENT APPLICATION NUMBER: US/09/782,816A  
CURRENT FILING DATE: 2001-02-14  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 17  
TYPE: PRT  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: A sequence conserved among Homo sapiens and Mus  
OTHER INFORMATION: musculus.  
US-09-782-816A-8

Query Match 35.1%; Score 88; DB 10; Length 17;  
Best Local Similarity 100.0%; Pred. No. 0.0048;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PQCKYQRLHVEVQLTT 22  
|||||  
Db 1 PQCKYQRLHVEVQLTT 17

Search completed: November 3, 2005, 22:11:40  
Job time : 93.7112 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2005, 21:32:33 ; Search time 72.5562 Seconds  
(without alignments)  
282.516 Million cell updates/sec

Title: 09782816-52

Sequence: 1 GKKTPTVQKCRLOIEMNEL.....ADERSQSDAVATVISTAR 53

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: geneseqp19908:\*  
2: geneseqp19908:\*  
3: geneseqp20008:\*  
4: geneseqp20008:\*  
5: geneseqp20028:\*  
6: geneseqp20038:\*  
7: geneseqp20038:\*  
8: geneseqp20048:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	258	100.0	53	5	ABP53017 Cellular
2	243.5	94.4	380	4	ABP53017 Cellular
3	243.5	94.4	380	5	ABP53020 Drosophila
4	114	44.2	22	5	ABP52991 Cellular
5	108	41.9	21	5	ABP52992 Cellular
6	105	40.7	22	5	ABP52993 Cellular
7	103	39.9	20	5	ABP52994 Cellular
8	98	38.0	19	5	ABP52995 Cellular
9	93	36.0	18	5	ABP52996 Cellular
10	88	34.1	17	5	ABP52997 Cellular
11	81	31.4	16	5	ABP52998 Cellular
12	77	29.8	15	5	ABP53019 Cellular
13	74	28.7	14	5	ABP53019 Cellular
14	72.5	28.1	14	5	ABP53016 Cellular
15	72.5	28.1	14	5	ABP53016 Cellular
16	72.5	28.1	14	5	ABP53016 Cellular
17	72.5	28.1	14	5	ABP53016 Cellular
18	72.5	28.1	14	5	ABP53016 Cellular
19	72.5	28.1	14	5	ABP53016 Cellular
20	72.5	28.1	14	5	ABP53016 Cellular
21	70.5	27.3	13	8	ADQ17289 Human sof
22	70.5	27.3	13	8	ADQ17242 Human sof
23	69.5	26.9	97	4	ABP38174 Peptide #
24	69.5	26.9	97	4	ABP38174 Peptide #
25	69.5	26.9	97	4	ABP38174 Peptide #

26	69.5	26.9	97	4	AAM58807 Human bra
27	69.5	26.9	97	4	ABG53031 Human liv
28	69.5	26.9	97	5	ABG41126 Human pep
29	67	26.0	13	5	ABP53000 Cellular
30	62.5	24.2	97	4	AAM19334 Peptide #
31	62.5	24.2	97	4	ABP38669 Peptide #
32	62.5	24.2	97	4	AAM32130 Peptide #
33	62.5	24.2	97	4	ABP23745 Protein #
34	62.5	24.2	97	4	AAM71845 Protein #
35	62.5	24.2	97	4	AAM59298 Human bra
36	62.5	24.2	97	4	ABG53529 Human liv
37	62.5	24.2	97	5	ABG41658 Human pep
38	62.5	24.2	103	4	ABP26895 Protein #
39	62.5	24.2	103	4	AAM77711 Human don
40	62.5	24.2	892	7	ADB79862 Rat myosi
41	62.5	24.2	892	7	ADD47859 Rat prote
42	62.5	24.2	892	7	ADD47855 Rat prote
43	62.5	24.2	1939	7	ADD47857 Human pro
44	62.5	24.2	1939	7	ADD47861 Human pro
45	62.5	24.2	1939	7	ADD47033 Human pro

## ALIGNMENTS

RESULT 1  
ABP53017  
ID ABP53017 standard; peptide; 53 AA.

AC ABP53017;

DT 05-NOV-2002 (first entry)

XX Cellular proliferation peptide inhibitor SEQ ID NO:52.

XX Cellular proliferation inhibition; cytostatic; antiinflammatory; cancer;

XX p50 inhibitor; dynamin inhibitor; gene therapy; tumour; carcinoma;

XX sarcoma; glioblastoma; leukaemia; lymphoid malignancy; neuronal disorder;

XX glial disorder; astrocytic disorder; hypothalamic disorder; inflammatory;

XX glandular disorder; macropagal disorder; epithelial disorder;

XX stromal disorder; blastocoele disorder; angiogenic disorder;

XX immunologic disorder.

XX Drosophila melanogaster.

OS WO200264779-A2.

XX 22-AUG-2002.

XX 21-JAN-2002; 2002WO-US001708.

XX 14-FEB-2001; 2001US-00782816.

XX (REGC ) UNIV CALIFORNIA.

XX Sharp DJ, Rogers GC, Scholey JW;

XX WPI; 2002-657599/70.

XX New peptide inhibitors of p50/dynamin useful for treating cancer by

XX inhibiting cellular proliferation, e.g. benign or malignant tumors,

XX leukemia and lymphoid malignancies, or inflammatory, angiogenic and

XX immunologic disorders.

XX Claim 3; Page 31; 55pp; English.

XX The present invention describes an isolated peptide (I) comprising or

XX having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the

XX sequences given in ABP52966 and ABP52967 and can have C-terminal and N-

XX terminal extensions. (I) have cytosolic and antiinflammatory activities

XX and can be used as p50/dynamin inhibitors and in gene therapy. The

XX peptides, nucleic acid molecules and methods from the present invention

XX are useful for treating cancer by inhibiting cellular proliferation, such

CC as benign or malignant tumours (renal, liver, kidney, bladder, breast,  
CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
CC neck tumours); leukemias and lymphoid malignancies, other disorders such  
CC as neuronal, glial, astrocytal, hypothalamic and other glandular,  
CC macrophagal, epithelial, stromal and blastocoealic disorders; and  
CC inflammatory, angiogenic and immunologic disorders. The present sequence  
CC represents a specifically claimed peptide inhibitor of cellular  
CC proliferation from the present invention

XX  
SQ Sequence 53 AA;

Query Match 100.0%; Score 258; DB 5; Length 53;

Best Local Similarity 100.0%; Pred. No. 2.4e-26; Mismatches 0; Indels 0; Gaps 0;

Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GKEKTPVQKCORLQIEMNELINEVAALQVDRKVADEEKOSYDAVAVTISTAR 53  
Db 1 GKEKTPVQKCORLQIEMNELINEVAALQVDRKVADEEKOSYDAVAVTISTAR 53

RESULT 2

ABBS9088 ID ABBS9088 standard; protein; 380 AA.

XX  
AC ABBS9088;

XX  
DT 26-MAR-2002 (first entry)

XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 4056.

XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
XX pharmaceutical.

XX  
OS Drosophila melanogaster.

XX  
PN WO200171042-A2.

XX  
XX 27-SEP-2001.

XX  
PD 23-MAR-2001; 2001WO-US0009231.

XX  
PP 23-MAR-2000; 2000US-0191637P.

XX  
PR 11-JUL-2000; 2000US-00614150.

XX  
XX (PEKE ) PE CORP NY.

XX  
PI Venter JC, Adams M, Li PWD, Myers EW;

XX  
DR WPI; 2001-656860/75.

XX  
DR N-PSDB; ABL03191.

XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signaling and cell-cell  
XX interactions.

XX  
PS Disclosure; SEQ ID NO 4056; 21pp + Sequence Listing; English.

XX  
CC The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
XX useful in developmental biology and in elucidating cell signalling and  
XX cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention  
XX discloses genomic DNA sequences (ABLU6176-ABLU30511), expressed DNA  
XX sequences (ABLU6140-ABLU6175) and the encoded proteins (ABBS7737-  
XX ABBS72072). The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 380 AA;

Query Match 94.4%; Score 243.5; DB 4; Length 380;  
Best Local Similarity 98.1%; Pred. No. 2.4e-23;

Matches 52; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 GKEKTPVQKCORLQIEMNELINEVAALQVDRKVADEEKOSYDAVAVTISTAR 53  
Db 94 GKEKTPVQKCORLQIEMNELINEVAALQVDRKVADEEKOSYDAVAVTISTAR 145

RESULT 3

ABPS3020 ID ABPS3020 standard; protein; 380 AA.

XX  
AC ABPS3020;

XX  
DT 05-NOV-2002 (first entry)

XX  
DE Drosophila melanogaster p50 protein sequence SEQ ID NO:56.

XX  
KW Cellular proliferation inhibition; cytostatic; antiinflammatory; cancer;  
KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;  
KW glandular disorder; macrophagal disorder; epithelial disorder;  
KW stromal disorder; blastocoealic disorder; angiogenic disorder;  
KW immunologic disorder; p50.

XX  
OS Drosophila melanogaster.

XX  
PN WO200264779-A2.

XX  
PD 22-AUG-2002.

XX  
PE 21-JAN-2002; 2002WO-US001708.

XX  
PR 14-FEB-2001; 2001US-00782816.

XX  
XX (REGC ) UNIV CALIFORNIA.

XX  
PI Sharp DJ, Rogers GC, Scholey JW;

XX  
DR WPI; 2002-657599/70.

XX  
DR N-PSDB; ABQ75378.

XX  
PT New peptide inhibitors of p50/dynaminin useful for treating cancer by  
XX inhibiting cellular proliferation, e.g. benign or malignant tumors,  
XX leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
XX immunologic disorders.

XX  
PS Example 1; Fig 4; 55pp; English.

XX  
CC The present invention describes an isolated peptide (I) comprising or  
XX having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
XX sequences given in ABPS2966 and ABPS2967 and can have C-terminal and N-  
XX terminal extensions. (I) have cytostatic and antiinflammatory activities  
XX and can be used as p50/dynaminin inhibitors and in gene therapy. The  
XX peptides, nucleic acid molecules and methods from the present invention  
XX are useful for treating cancer by inhibiting cellular proliferation, such  
XX as benign or malignant tumours (renal, liver, kidney, bladder, breast,  
XX gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
XX thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
XX neck tumours); leukemias and lymphoid malignancies, other disorders such  
XX as neuronal, glial, astrocytal, hypothalamic and other glandular,  
XX macrophagal, epithelial, stromal and blastocoealic disorders; and  
XX inflammatory, angiogenic and immunologic disorders. The present sequence  
XX represents Drosophila melanogaster p50 which is given in the  
XX exemplification of the present invention

SQ Sequence 380 AA;

Query Match 94.4%; Score 243.5; DB 5; Length 380;  
Best Local Similarity 98.1%; Pred. No. 2.4e-23; Mismatches 0; Indels 1; Gaps 1;  
Matches 52; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 GKEKTPVQKCORLQIEMNELINEVAALQVDRKVADEEKOSYDAVAVTISTAR 53

Db 94 GSEKTPVOKCQRLQIEMNELINVAALQVDRKVADEEKQSYDA-VAIVISTAR 145.

## RESULT 4

ABP52991 ID ABP52991 standard; peptide: 22 AA.

AC ABP52991;

DT 05-NOV-2002 (first entry)

DE Cellular proliferation inhibitor related peptide SEQ ID NO:26.

XX Cellular proliferation inhibition; cytostatic; antiinflammatory; cancer;  
KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytoma; astrocytic disorder; hypothalamic disorder; inflammatory;  
KW glandular disorder; macrophagal disorder; epithelial disorder;  
KW stromal disorder; blastocoele disorder; angiogenic disorder;  
KW immunologic disorder.

OS Drosophila melanogaster.

PN WO200264779-A2.

PD 22-AUG-2002.

PE 21-JAN-2002; 2002WO-US001708.

PR 14-FEB-2001; 2001US-00782816.

PA (REGC ) UNIV CALIFORNIA.

PI Sharp DJ, Rogers GC, Scholey JM;

PT WPI; 2002-657599/70.

XX New peptide inhibitors of p50/dynaminin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.

PS Claim 1; Page 30; 55pp; English.

XX The present invention describes an isolated peptide (I) comprising or  
CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
CC terminal extensions. (I) have cytostatic and antiinflammatory activities  
CC and can be used as p50/dynaminin inhibitors and in gene therapy. The  
CC peptides, nucleic acid molecules and methods from the present invention  
CC are useful for treating cancer by inhibiting cellular proliferation, such  
CC as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
CC neck tumors); leukemias and lymphoid malignancies, other disorders such  
CC as neuronal, glial, astrocytic, hypothalamic and other glandular,  
CC macrophagal, epithelial, stromal and blastocoele disorders; and  
CC inflammatory, angiogenic and immunologic disorders. The present sequence  
CC represents a peptide that can be N-terminally added to (P2)

XX Sequence 22 AA;

Query Match 44.2%; Score 114; DB 5; Length 22;

Best Local Similarity 100.0%; Pred. No. 8e-08; Mismatches 0; Indels 0; Gaps 0;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSEKTPVOKCQRLQIEMNELIN 22

Db 1 GSEKTPVOKCQRLQIEMNELIN 22

RESULT 5

ABP52992 ID ABP52992 standard; peptide: 21 AA.

AC ABP52992;

DT 05-NOV-2002 (first entry)

DE Cellular proliferation inhibitor related peptide SEQ ID NO:27.

XX Cellular proliferation inhibition; cytostatic; antiinflammatory; cancer;  
KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytoma; astrocytic disorder; hypothalamic disorder; inflammatory;  
KW glandular disorder; macrophagal disorder; epithelial disorder;  
KW stromal disorder; blastocoele disorder; angiogenic disorder;  
KW immunologic disorder.

OS Drosophila melanogaster.

PN WO200264779-A2.

PD 22-AUG-2002.

PE 21-JAN-2002; 2002WO-US001708.

PR 14-FEB-2001; 2001US-00782816.

PA (REGC ) UNIV CALIFORNIA.

PI Sharp DJ, Rogers GC, Scholey JM;

PT WPI; 2002-657599/70.

XX New peptide inhibitors of p50/dynaminin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.

PS Claim 1; Page 30; 55pp; English.

XX The present invention describes an isolated peptide (I) comprising or  
CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
CC terminal extensions. (I) have cytostatic and antiinflammatory activities  
CC and can be used as p50/dynaminin inhibitors and in gene therapy. The  
CC peptides, nucleic acid molecules and methods from the present invention  
CC are useful for treating cancer by inhibiting cellular proliferation, such  
CC as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
CC neck tumors); leukemias and lymphoid malignancies, other disorders such  
CC as neuronal, glial, astrocytic, hypothalamic and other glandular,  
CC macrophagal, epithelial, stromal and blastocoele disorders; and  
CC inflammatory, angiogenic and immunologic disorders. The present sequence  
CC represents a peptide that can be N-terminally added to (P2)

XX Sequence 21 AA;

Query Match 41.9%; Score 108; DB 5; Length 21;

Best Local Similarity 100.0%; Pred. No. 4.7e-07; Mismatches 0; Indels 0; Gaps 0;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 EKETPVOKCQRLQIEMNELIN 22

Db 1 EKETPVOKCQRLQIEMNELIN 21

## RESULT 6

ABP52967 ID ABP52967 standard; peptide: 22 AA.

AC ABP52967;

XX

DT 05-NOV-2002 (first entry)  
 XX Cellular proliferation inhibitor related peptide SEQ ID NO:2.  
 XX  
 XX Cellular proliferation inhibitor; cytostatic; antiinflammatory; cancer;  
 KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
 KW sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
 KW glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;  
 KW glandular disorder; macropagal disorder; epithelial disorder;  
 KW stromal disorder; blastocoeleic disorder; angiogenic disorder;  
 KW immunologic disorder.  
 KW  
 XX Drosophila melanogaster.  
 OS  
 XX WO200264779-A2.  
 PN  
 XX 22-AUG-2002.  
 PD  
 XX 21-JAN-2002; 2002WO-US001708.  
 PP  
 XX 14-FEB-2001; 2001US-00782816.  
 PR  
 XX (REGC ) UNIV CALIFORNIA.  
 PA  
 XX Sharp DJ, Rogers GC, Scholey JM;  
 PI  
 XX WPI; 2002-657599/70.  
 DR  
 XX  
 XX New peptide inhibitors of p50/dynaminin useful for treating cancer by  
 PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
 PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
 PT immunologic disorders.  
 PT  
 XX Claim 1; Page 29; 55pp; English.  
 PS  
 XX The present invention describes an isolated peptide (I) comprising or  
 CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
 CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
 CC terminal extensions. (I) have cytostatic and antiinflammatory activities  
 CC and can be used as p50/dynaminin inhibitors and in gene therapy. The  
 CC peptides, nucleic acid molecules and methods from the present invention  
 CC are useful for treating cancer by inhibiting cellular proliferation, such  
 CC as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
 CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
 CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
 CC neck tumors); leukemias and lymphoid malignancies, other disorders such  
 CC as neuronal, glial, astrocytal, hypothalamic and other glandular,  
 CC macropagal, epithelial, stromal and blastocoeleic disorders; and  
 CC inflammatory, angiogenic and immunologic disorders  
 CC  
 XX Sequence 22 AA;  
 SQ  
 Query Match 40.7%; Score 105; DB 5; Length 22;  
 Best Local Similarity 100.0%; Pred. NO. 1.2e-06;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 23 EVAALQVDRKVADEKQSYDAY 44  
 DB 1 EVAALQVDRKVADEKQSYDAY 22  
 RESULT 7  
 ABP52993  
 ID ABP52993 standard; peptide; 20 AA.  
 AC ABP52993;  
 XX  
 DT 05-NOV-2002 (first entry)  
 XX  
 DE Cellular proliferation inhibitor related peptide SEQ ID NO:28.  
 KW Cellular proliferation inhibitor; cytostatic; antiinflammatory; cancer;  
 KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
 KW immunologic disorder.

KW sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
 KW glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;  
 KW glandular disorder; macropagal disorder; epithelial disorder;  
 KW stromal disorder; blastocoeleic disorder; angiogenic disorder;  
 KW immunologic disorder.  
 KW  
 XX Drosophila melanogaster.  
 OS  
 XX WO200264779-A2.  
 PN  
 XX 22-AUG-2002.  
 PD  
 XX 21-JAN-2002; 2002WO-US001708.  
 PP  
 XX 14-FEB-2001; 2001US-00782816.  
 PR  
 XX (REGC ) UNIV CALIFORNIA.  
 PA  
 XX Sharp DJ, Rogers GC, Scholey JM;  
 PI  
 XX WPI; 2002-657599/70.  
 DR  
 XX  
 XX New peptide inhibitors of p50/dynaminin useful for treating cancer by  
 PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
 PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
 PT immunologic disorders.  
 PT  
 XX Claim 1; Page 30; 55pp; English.  
 PS  
 XX The present invention describes an isolated peptide (I) comprising or  
 CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
 CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
 CC terminal extensions. (I) have cytostatic and antiinflammatory activities  
 CC and can be used as p50/dynaminin inhibitors and in gene therapy. The  
 CC peptides, nucleic acid molecules and methods from the present invention  
 CC are useful for treating cancer by inhibiting cellular proliferation, such  
 CC as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
 CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
 CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
 CC neck tumors); leukemias and lymphoid malignancies, other disorders such  
 CC as neuronal, glial, astrocytal, hypothalamic and other glandular,  
 CC macropagal, epithelial, stromal and blastocoeleic disorders; and  
 CC inflammatory, angiogenic and immunologic disorders. The present sequence  
 CC represents a peptide that can be N-terminally added to (P2)  
 CC  
 XX Sequence 20 AA;  
 SQ  
 Query Match 39.9%; Score 103; DB 5; Length 20;  
 Best Local Similarity 100.0%; Pred. NO. 2e-06;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 KETPVOKCQRQIEMNELLN 22  
 DB 1 KETPVOKCQRQIEMNELLN 20  
 RESULT 8  
 ABP52994  
 ID ABP52994 standard; peptide; 19 AA.  
 AC ABP52994;  
 XX  
 DT 05-NOV-2002 (first entry)  
 XX  
 DE Cellular proliferation inhibitor related peptide SEQ ID NO:29.  
 KW Cellular proliferation inhibitor; cytostatic; antiinflammatory; cancer;  
 KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
 KW sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
 KW glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;  
 KW glandular disorder; macropagal disorder; epithelial disorder;  
 KW stromal disorder; blastocoeleic disorder; angiogenic disorder;  
 KW immunologic disorder.

```

XX OS Drosophila melanogaster.
XX PN WO200264779-A2.
XX PD 22-AUG-2002.
XX PF 21-JAN-2002; 2002WO-US001708.
XX PR 14-FEB-2001; 2001US-00782816.
XX PA (REGC ) UNIV CALIFORNIA.
XX PI Sharp DJ, Rogers GC, Scholey JM;
XX DR WPI; 2002-657599/70.
XX PT New peptide inhibitors of p50/dynaminin useful for treating cancer by
XX PT inhibiting cellular proliferation, e.g. benign or malignant tumors,
XX PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and
XX PT immunologic disorders.
XX PS Claim 1; Page 30; 55pp; English.
XX CC The present invention describes an isolated peptide (1) comprising or
XX CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the
XX CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-
XX CC terminal extensions. (1) have cytosstatic and antiinflammatory activities
XX CC and can be used as p50/dynaminin inhibitors and in gene therapy. The
XX CC peptides, nucleic acid molecules and methods from the present invention
XX CC are useful for treating cancer by inhibiting cellular proliferation, such
XX CC as benign or malignant tumors (renal, liver, kidney, bladder, breast,
XX CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,
XX CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, other disorders such
XX CC as neuronal, glial, astrocytal, hypothalamic and other glandular,
XX CC macrophagal, epithelial, stromal and blastocoeleic disorders; and
XX CC inflammatory, angiogenic and immunologic disorders. The present sequence
XX CC represents a peptide that can be N-terminally added to (P2)
XX SO
Sequence 19 AA;
Query Match 38.0%; Score 98; DB 5; Length 19;
Best Local Similarity 100.0%; Pred. No. 8.6e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 ETPVOKCORLOIEMNELIN 22
DB 1 ETPVOKCORLOIEMNELIN 19
RESULT 9
ABP52995
ID ABP52995 standard; peptide; 18 AA.
XX AC ABP52995;
XX DT 05-NOV-2002 (first entry)
XX DE Cellular proliferation inhibitor related peptide SEQ ID NO:30.
XX KW Cellular proliferation inhibition; cytosstatic; antiinflammatory; cancer;
XX KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;
XX KW sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;
XX KW glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;
XX KW glandular disorder; macrophagal disorder; epithelial disorder;
XX KW stromal disorder; blastocoeleic disorder; angiogenic disorder;
XX KW immunologic disorder.
XX OS Drosophila melanogaster.
XX PN WO200264779-A2.

```

```

PD 22-AUG-2002.
XX PF 21-JAN-2002; 2002WO-US001708.
XX PR 14-FEB-2001; 2001US-00782816.
XX PA (REGC ) UNIV CALIFORNIA.
XX PI Sharp DJ, Rogers GC, Scholey JM;
XX DR WPI; 2002-657599/70.
XX PT New peptide inhibitors of p50/dynaminin useful for treating cancer by
XX PT inhibiting cellular proliferation, e.g. benign or malignant tumors,
XX PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and
XX PT immunologic disorders.
XX PS Claim 1; Page 30; 55pp; English.
XX CC The present invention describes an isolated peptide (1) comprising or
XX CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the
XX CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-
XX CC terminal extensions. (1) have cytosstatic and antiinflammatory activities
XX CC and can be used as p50/dynaminin inhibitors and in gene therapy. The
XX CC peptides, nucleic acid molecules and methods from the present invention
XX CC are useful for treating cancer by inhibiting cellular proliferation, such
XX CC as benign or malignant tumors (renal, liver, kidney, bladder, breast,
XX CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,
XX CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, other disorders such
XX CC as neuronal, glial, astrocytal, hypothalamic and other glandular,
XX CC macrophagal, epithelial, stromal and blastocoeleic disorders; and
XX CC inflammatory, angiogenic and immunologic disorders. The present sequence
XX CC represents a peptide that can be N-terminally added to (P2)
XX SO
Sequence 18 AA;
Query Match 36.0%; Score 93; DB 5; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.7e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 TPVOKCORLOIEMNELIN 22
DB 1 TPVOKCORLOIEMNELIN 18
RESULT 10
ABP52996
ID ABP52996 standard; peptide; 17 AA.
XX AC ABP52996;
XX DT 05-NOV-2002 (first entry)
XX DE Cellular proliferation inhibitor related peptide SEQ ID NO:31.
XX KW Cellular proliferation inhibition; cytosstatic; antiinflammatory; cancer;
XX KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;
XX KW sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;
XX KW glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;
XX KW glandular disorder; macrophagal disorder; epithelial disorder;
XX KW stromal disorder; blastocoeleic disorder; angiogenic disorder;
XX KW immunologic disorder.
XX OS Drosophila melanogaster.
XX PN WO200264779-A2.
XX PD 22-AUG-2002.
XX PF 21-JAN-2002; 2002WO-US001708.
XX PR 14-FEB-2001; 2001US-00782816.

```

XX (REGC ) UNIV CALIFORNIA.  
PA Sharp DJ, Rogers GC, Scholey JM;  
XX WPI, 2002-657599/70.  
XX  
XX New peptide inhibitors of p50/dynactin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.  
XX  
XX Claim 1; Page 30; 55pp; English.  
XX  
XX The present invention describes an isolated peptide (1) comprising or  
CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
CC terminal extensions. (1) have cytosstatic and antiinflammatory activities  
CC and can be used as p50/dynactin inhibitors and in gene therapy. The  
CC peptides, nucleic acid molecules and methods from the present invention  
CC are useful for treating cancer by inhibiting cellular proliferation, such  
CC as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
CC neck tumors); leukemias and lymphoid malignancies, other disorders such  
CC as neuronal, glial, astrocytal, hypothalamic and other glandular,  
CC macrophagal, epithelial, stromal and blastocoealic disorders; and  
CC inflammatory, angiogenic and immunologic disorders. The present sequence  
CC represents a peptide that can be N-terminally added to (P2)  
XX  
SQ Sequence 17 AA;  
XX  
Query Match 34.1%; Score 88; DB 5; Length 17;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 6 PVQKCRQLQIEMNELLN 22  
Db 1 PVQKCRQLQIEMNELLN 17  
XX  
RESULT 11  
ABP52997  
ID ABP52997 standard; peptide; 16 AA.  
XX  
AC ABP52997;  
XX  
DT 05-NOV-2002 (first entry)  
XX  
DE Cellular proliferation inhibitor related peptide SEQ ID NO:32.  
XX  
KW Cellular proliferation inhibition; cytosstatic; antiinflammatory; cancer;  
KW p50 inhibitor; dynactin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;  
KW glandular disorder; macrophagal disorder; epithelial disorder;  
KW stromal disorder; blastocoealic disorder; angiogenic disorder;  
KW immunologic disorder.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200264779-A2.  
XX  
PD 22-AUG-2002.  
XX  
PF 21-JAN-2002; 2002WO-US001708.  
XX  
PR 14-FEB-2001; 2001US-00782816.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX Sharp DJ, Rogers GC, Scholey JM;  
XX  
XX

DR WPI, 2002-657599/70.  
XX  
XX New peptide inhibitors of p50/dynactin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.  
XX  
XX Claim 1; Page 30; 55pp; English.  
XX  
XX The present invention describes an isolated peptide (1) comprising or  
CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
CC terminal extensions. (1) have cytosstatic and antiinflammatory activities  
CC and can be used as p50/dynactin inhibitors and in gene therapy. The  
CC peptides, nucleic acid molecules and methods from the present invention  
CC are useful for treating cancer by inhibiting cellular proliferation, such  
CC as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
CC neck tumors); leukemias and lymphoid malignancies, other disorders such  
CC as neuronal, glial, astrocytal, hypothalamic and other glandular,  
CC macrophagal, epithelial, stromal and blastocoealic disorders; and  
CC inflammatory, angiogenic and immunologic disorders. The present sequence  
CC represents a peptide that can be N-terminally added to (P2)  
XX  
SQ Sequence 16 AA;  
XX  
Query Match 31.4%; Score 81; DB 5; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.0012;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 7 VQKCRQLQIEMNELLN 22  
Db 1 VQKCRQLQIEMNELLN 16  
XX  
RESULT 12  
ABP52998  
ID ABP52998 standard; peptide; 15 AA.  
XX  
AC ABP52998;  
XX  
DT 05-NOV-2002 (first entry)  
XX  
DE Cellular proliferation inhibitor related peptide SEQ ID NO:33.  
XX  
KW Cellular proliferation inhibition; cytosstatic; antiinflammatory; cancer;  
KW p50 inhibitor; dynactin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;  
KW glandular disorder; macrophagal disorder; epithelial disorder;  
KW stromal disorder; blastocoealic disorder; angiogenic disorder;  
KW immunologic disorder.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200264779-A2.  
XX  
PD 22-AUG-2002.  
XX  
PF 21-JAN-2002; 2002WO-US001708.  
XX  
PR 14-FEB-2001; 2001US-00782816.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX Sharp DJ, Rogers GC, Scholey JM;  
XX  
XX WPI, 2002-657599/70.  
XX  
XX New peptide inhibitors of p50/dynactin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and

PT immunologic disorders.  
 XX  
 PS Claim 1, Page 30, 55pp; English.  
 XX  
 CC The present invention describes an isolated peptide (I) comprising or  
 CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
 CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
 CC terminal extensions. (1) have cytostatic and anti-inflammatory activities  
 CC and can be used as p50/dynaminin inhibitors and in gene therapy. The  
 CC peptides, nucleic acid molecules and methods from the present invention  
 CC are useful for treating cancer by inhibiting cellular proliferation, such  
 CC as benign or malignant tumours (renal, liver, kidney, bladder, breast,  
 CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
 CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
 CC neck tumours); leukaemias and lymphoid malignancies, other disorders such  
 CC as neuronal, glial, astrocytal, hypochalamic and other glandular,  
 CC macrophagal, epithelial, stromal and blastocoeleic disorders; and  
 CC inflammatory, angiogenic and immunologic disorders. The present sequence  
 CC represents a peptide that can be N-terminally added to (P2)  
 XX  
 SQ Sequence 15 AA;  
 Query Match 29.8%; Score 77; DB 5; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.0038;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 8 QKQRLQIEMNELIN 22  
 |||||  
 1 QKQRLQIEMNELIN 15  
 DB  
 RESULT 13  
 ABP53019  
 ID ABP53019 standard; protein; 183 AA.  
 XX  
 AC ABP53019;  
 XX  
 DT 05-NOV-2002 (first entry)  
 XX  
 DE Mouse p50 amino acid sequence SEQ ID NO:54.  
 XX  
 KW Cellular proliferation inhibition; cytostatic; anti-inflammatory; cancer;  
 KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
 KW sarcoma; glioblastoma; leukaemia; lymphoid malignancy; neuronal disorder;  
 KW glial disorder; astrocytal disorder; hypochalamic disorder; inflammatory;  
 KW glandular disorder; macrophagal disorder; epithelial disorder;  
 KW stromal disorder; blastocoeleic disorder; angiogenic disorder;  
 KW immunologic disorder.  
 XX  
 OS Mus musculus.  
 XX  
 PN WO200264779-A2.  
 XX  
 PD 22-AUG-2002.  
 XX  
 PF 21-JAN-2002; 2002WO-US001708.  
 XX  
 PR 14-FEB-2001; 2001US-00782816.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Sharp DJ, Rogers GC, Scholey JM,  
 DR WPI; 2002-657599/70.  
 XX  
 XX New peptide inhibitors of p50/dynaminin useful for treating cancer by  
 PT inhibiting cellular proliferation, e.g. benign or malignant tumours,  
 PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
 PT immunologic disorders.  
 XX  
 PS Disclosure; Fig 2; 55pp; English.  
 XX  
 CC The present invention describes an isolated peptide (I) comprising or

CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
 CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
 CC terminal extensions. (1) have cytostatic and anti-inflammatory activities  
 CC and can be used as p50/dynaminin inhibitors and in gene therapy. The  
 CC peptides, nucleic acid molecules and methods from the present invention  
 CC are useful for treating cancer by inhibiting cellular proliferation, such  
 CC as benign or malignant tumours (renal, liver, kidney, bladder, breast,  
 CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
 CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
 CC neck tumours); leukaemias and lymphoid malignancies, other disorders such  
 CC as neuronal, glial, astrocytal, hypochalamic and other glandular,  
 CC macrophagal, epithelial, stromal and blastocoeleic disorders; and  
 CC inflammatory, angiogenic and immunologic disorders. The present sequence  
 CC represents mouse p50 which is given in the exemplification of the present  
 CC invention  
 XX  
 SQ Sequence 183 AA;  
 Query Match 28.7%; Score 74; DB 5; Length 183;  
 Best Local Similarity 35.8%; Pred. No. 0.23;  
 Matches 19; Conservative 10; Mismatches 24; Indels 0; Gaps 0;  
 Oy 1 GKEETPVQKQRLQIEMNELINEVALQVDRKVADEEKOSYDAVATVISTAR 53  
 |||||  
 94 GVKETPQOKTQRLQIEMNELINEVALQVDRKVADEEKOSYDAVATVISTAR 146  
 DB  
 RESULT 14  
 ABP53016  
 ID ABP53016 standard; peptide; 52 AA.  
 XX  
 AC ABP53016;  
 XX  
 DT 05-NOV-2002 (first entry)  
 XX  
 DE Cellular proliferation peptide inhibitor SEQ ID NO:51.  
 XX  
 KW Cellular proliferation inhibition; cytostatic; anti-inflammatory; cancer;  
 KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
 KW sarcoma; glioblastoma; leukaemia; lymphoid malignancy; neuronal disorder;  
 KW glial disorder; astrocytal disorder; hypochalamic disorder; inflammatory;  
 KW glandular disorder; macrophagal disorder; epithelial disorder;  
 KW stromal disorder; blastocoeleic disorder; angiogenic disorder;  
 KW immunologic disorder.  
 XX  
 OS Homo sapiens.  
 OS Mus musculus.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 44  
 FT /label= 'Leu, Val  
 FT /note= "Leu in humans and Val in Mus musculus"  
 XX  
 PN WO200264779-A2.  
 XX  
 PD 22-AUG-2002.  
 XX  
 PF 21-JAN-2002; 2002WO-US001708.  
 XX  
 PR 14-FEB-2001; 2001US-00782816.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Sharp DJ, Rogers GC, Scholey JM,  
 DR WPI; 2002-657599/70.  
 XX  
 XX New peptide inhibitors of p50/dynaminin useful for treating cancer by  
 PT inhibiting cellular proliferation, e.g. benign or malignant tumours,  
 PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
 PT immunologic disorders.  
 XX

PS Claim 2; Page 31; 55pp; English.  
XX  
CC The present invention describes an isolated peptide (I) comprising or  
CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
CC terminal extensions. (1) Have cytostatic and anti-inflammatory activities  
CC and can be used as p50/dynactin inhibitors and in gene therapy. The  
CC peptides, nucleic acid molecules and methods from the present invention  
CC are useful for treating cancer by inhibiting cellular proliferation, such  
CC as benign or malignant tumours (renal, liver, kidney, bladder, breast,  
CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
CC neck tumours); leukemias and lymphoid malignancies, other disorders such  
CC as neuronal, glial, astrocytal, hypochalamic and other glandular,  
CC macrophagal, epithelial, stromal and blastocoelec disorders; and  
CC inflammatory, angiogenic and immunologic disorders. The present sequence  
CC represents a specifically claimed peptide inhibitor of cellular  
CC proliferation from the present invention  
SQ Sequence 52 AA;  
Query Match 28.1%; Score 72.5; DB 5; Length 52;  
Best Local Similarity 48.7%; Pred. No. 0.072; Mismatches 14; Indels 1; Gaps 1;  
Matches 19; Conservative 5; Mismatches 14; Indels 1; Gaps 1;  
OY 1 GKEKTPVQKCORLQIEMNELLNEVALQVD-RKYADEEK 38  
DB 1 GVKETPQOKYORLHVEOELTTEVEKIKTYKESATEEK 39  
RESULT 15  
ABM81421  
ID ABM81421 standard; protein; 314 AA.  
AC ABM81421;  
DT 18-NOV-2004 (first entry)  
XX  
DE Tumour-associated antigenic target (TAT) polypeptide PRO82191, SEQ:3675.  
XX  
KM Tumour-associated antigenic target; TAT; human; overexpression; cancer;  
KM tumour; diagnosis; cell proliferative disorder; breast cancer;  
KM colorectal cancer; lung cancer; ovarian cancer; liver cancer;  
KM central nervous system cancer; bladder cancer; pancreatic cancer;  
KM cervical cancer; melanoma; leukaemia; hybridisation probe;  
KM chromosome identification; chromosome mapping; gene mapping;  
KM gene therapy; cytostatic.  
XX  
OS Homo sapiens.  
XX  
PN WO2004030615-A2.  
XX  
PD 15-APR-2004.  
XX  
PF 29-SEP-2003; 2003WO-US028547.  
XX  
PR 02-OCT-2002; 2002US-0414971P.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Wu TD, Zhang Z, Zhou Y;  
XX  
DR WPI: 2004-347921/32.  
DR N-PSDB; ACN39497.  
XX  
PT New tumor-associated antigenic target polypeptides and nucleic acids,  
PT useful in preparing a medicament for treating or detecting a  
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or  
PT prostate cancer or tumor.  
XX  
PS Claim 12; SEQ ID NO 3675; 7273pp; English.  
XX  
CC The invention relates to human tumour-associated antigenic target (TAT)

CC polypeptides, and their related nucleic acids. The TAT polypeptides are  
CC overexpressed in cancer tissues compared to normal tissues, and may thus  
CC serve as effective targets for the diagnosis and treatment of cancer in  
CC mammals. The invention also relates to nucleic acid and polypeptide  
CC sequences at least 80% identical to the TAT nucleic acids and  
CC polypeptides; expression vectors and host cells comprising a TAT nucleic  
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic  
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a  
CC TAT polypeptide; and methods and compositions for the treatment or  
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,  
CC antibodies, antagonists, binding molecules and compositions are useful  
CC for diagnosing or treating a cell proliferative disorder associated with  
CC increased TAT expression, particularly cancers such as breast cancer,  
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder  
CC cancer, pancreatic cancer, cervical cancer, cancers of the central  
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be  
CC used as hybridisation probes, in chromosome and gene mapping, in  
CC chromosome identification and in gene therapy. The present sequence  
CC represents a TAT polypeptide of the invention  
SQ Sequence 314 AA;  
Query Match 28.1%; Score 72.5; DB 8; Length 314;  
Best Local Similarity 48.7%; Pred. No. 0.71; Mismatches 14; Indels 1; Gaps 1;  
Matches 19; Conservative 5; Mismatches 14; Indels 1; Gaps 1;  
OY 1 GKEKTPVQKCORLQIEMNELLNEVALQVD-RKYADEEK 38  
DB 7 GVKETPQOKYORLHVEOELTTEVEKIKTYKESATEEK 45

Search completed: November 3, 2005, 21:57:31  
Job time : 73.5562 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2005, 21:43:09 ; Search time 19.2727 Seconds  
(without alignments)  
264.596 Million cell updates/sec

Title: 09782816-52

Sequence: 1 GKEETPVQKCRQLQIEMNEL.....ADEEKSYDAVAVTISTAR 53

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70.5	27.3	1937	2	I38055 myosin heavy chain
2	69.5	26.9	230	2	AC1523 hypothetical prote
3	68	26.4	297	2	H72670 hypothetical prote
4	67.5	26.2	162	2	H70314 hypothetical prote
5	66	25.6	508	2	C84739 hypothetical prote
6	65	25.2	730	2	T38346 hypothetical prote
7	64.5	25.0	157	2	B81688 conserved hypotnet
8	63.5	24.6	676	2	S00084 myosin heavy chain
9	63.5	24.6	698	2	S52696 myosin heavy chain
10	63.5	24.6	1938	1	UX0178 myosin heavy chain
11	63	24.4	111	2	F64596 hypothetical prote
12	62.5	24.2	222	2	B84283 TRK potassium upda
13	62.5	24.2	955	2	S24348 myosin heavy chain
14	62.5	24.2	1938	2	A59293 skeletal myosin he
15	62	24.0	230	2	AC1164 hypothetical prote
16	62	24.0	596	2	H81095 conserved hypotnet
17	62	24.0	701	2	F81846 hypothetical secre
18	62	24.0	743	2	D84854 hypothetical prote
19	62	24.0	1133	2	T22976 hypothetical prote
20	60.5	23.4	900	2	T33482 hypothetical prote
21	60	23.3	1050	2	A89769 hypothetical prote
22	59.5	23.1	534	2	A59088 ABC transporter -
23	59.5	23.1	741	2	S39082 myosin heavy chain
24	59.5	23.1	936	2	S39083 myosin heavy chain
25	59.5	23.1	1940	1	S04090 myosin heavy chain
26	59	22.9	110	2	H71915 phosphoglycerate m
27	58.5	22.7	503	2	A82831 N utilization subs
28	58.5	22.7	1940	1	A24922 myosin heavy chain
29	58.5	22.7	1940	1	A24922 myosin heavy chain

30	58	22.5	223	2	P00514 hemagglutinin HA2
31	58	22.5	412	2	S07537 myosin heavy chain
32	58	22.5	463	2	G97884 argininosuccinate
33	58	22.5	621	2	S10450 myosin heavy chain
34	58	22.5	650	2	T00617 endosyle-specific
35	58	22.5	866	2	S04027 paramyosin - Caeno
36	58	22.5	872	2	T19296 hypothetical prote
37	58	22.5	1255	2	T31065 diaphanous protein
38	58	22.5	2297	2	A82494 hypothetical prote
39	57.5	22.3	271	2	A30592 giardin beta chain
40	57.5	22.3	272	1	SJG086 hypothetical prote
41	57.5	22.3	410	2	H95888 protein YKR029c ho
42	57.5	22.3	560	2	S53382 hypothetical prote
43	57.5	22.3	629	2	T34370 hypothetical prote
44	57.5	22.3	791	2	T27473 myosin heavy chain
45	57.5	22.3	876	2	A23767 myosin heavy chain

#### ALIGNMENTS

RESULT 1  
I38055  
myosin heavy chain, perinatal skeletal muscle - human  
N:Contains: myosin ATPase (RC 3.6.4.1)  
C:Species: Homo sapiens (man)  
C>Date: 17-May-1996 #sequence revision 17-May-1996 #text change 09-Jul-2004  
C/Accession: I38055; JH0154; S12459; S09332; A30220; S49478  
R:Jullian, B.H.; Kelly, A.M.; Pompidou, A.J.; Hoffman, R.; Schiaffino, S.; Steedman, H.H.  
Eur. J. Biochem. 230, 1001-1006, 1995  
A>Title: Characterization of a human perinatal myosin heavy-chain transcript.  
A/Reference number: JH0154; MUID:90323631; PMID:2373371  
A/Accession: I38055; MUID:95324556; PMID:1601129  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1937 <RES>  
A/Cross-references: UNIPROT:P13535; EMBL:Z38133; NID:G558668; PIDD:CA86293.1; PID:G5586  
R:Karsch-Mizrachi, I.; Peghali, R.; Shows, T.B.; Leinwand, L.A.  
Gene 89, 289-294, 1990  
A>Title: Generation of a full-length human perinatal myosin heavy-chain-encoding cDNA.  
A/Reference number: JH0154; MUID:90323631; PMID:2373371  
A/Accession: JH0154  
A:Molecule type: mRNA  
A:Residues: 1-14, 'A', 16-85 <KAR>  
A/Cross-references: GB:Y00821  
A/Experimental source: skeletal muscle  
R:Bober, E.  
submitted to the EMBL Data Library, January 1989  
A/Reference number: S12458  
A/Accession: S12459  
A:Molecule type: mRNA  
A:Residues: 502-1071, 'N', 1073-1250, 'DGG', 1253-1376, 'NT', 1379-1913, 'D', 1915-1937 <BOB>  
A/Cross-references: EMBL:X51592; NID:929465; PIDD:CA83941.1; PID:G29466  
A/Experimental source: clone gEMHC-F  
R:Bober, E.; Buchberger-Seidl, A.; Braun, T.; Singh, S.; Goedde, H.W.; Arnold, H.H.  
Eur. J. Biochem. 189, 55-65, 1990  
A>Title: Identification of three developmentally controlled isoforms of human myosin hea  
A/Reference number: S09331; MUID:90235862; PMID:1691980  
A/Accession: S09332  
A:Molecule type: mRNA  
A:Residues: 502-547, 'X', 549-617, 'X', 619-687, 'X', 689-757, 'X', 759-827, 'X', 829-897, 'X', 899-  
1376, 'NT', 1379-1386, 'X', 1388-1456, 'X', 1458-1526, 'X', 1528-1596, 'X', 1598-1666, 'X', 1668-17  
A/Cross-references: EMBL:X51592  
R:Peghali, R.; Leinwand, L.A.  
J. Cell Biol. 108, 1791-1797, 1989  
A>Title: Molecular genetic characterization of a developmentally regulated human perinat  
A/Reference number: A30220; MUID:89234168; PMID:2715175  
A/Accession: A30220  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 860-969, 'Q', 971-1246, 'H', 1248-1260, 'G', 1262-1296, 'Q', 1298-1503, 'AH', 1506-184  
A/Cross-references: GB:Y00821; NID:G34863; PIDD:CA868757.1; PID:G34864  
C/Genetic8:

A:Gene: GDB:MYH8  
A:Cross-references: GDB:125267; OMIM:160741  
A:Map position: 17pter-17p12  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
C:Keywords: actin binding; ATP; coiled coil; hydrolase; muscle contraction; nucleotide b  
F:91-769/Domain: myosin motor domain homology <MMOT>  
F:181-188/Region: nucleotide-binding motif A (P-loop)  
F:551-588/Region: actin binding #status predicted  
F:658-680/Region: actin binding #status predicted  
F:847-1282/Region: S2 #status predicted  
F:698,708/Active site: Cys #status predicted

Query Match 27.3%; Score 70.5; DB 2; Length 1937;  
Best Local Similarity 28.9%; Pred. No. 12;  
Matches 22; Conservative 15; Mismatches 8; Indels 31; Gaps 4;

Oy 2 EKETPVQKQ-----RLQIEMNELINEVALQVDRKVA--DEE-- 37  
Db 1533 KQVOEOKCEIOALAEAEASLEHBEKILRIQLEINQKSB----VDKRIAEKDEID 1587

Oy 38 --KOSYDAVAVTVIST 51  
Db 1568 QLKRNHRTVETMST 1603

RESULT 2  
AC1523  
hypothetical protein lin0723 [imported] - *Listeria innocua* (strain C1p11262)  
C:Species: *Listeria innocua*  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AC1523  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioecker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fehli, H.  
D.; Jones, L.M.; Kars, U.  
Science 294, 849-852, 2001  
A:Authors: Krell, U.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,  
A:Title: Comparative genome of *Listeria species*.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AC1523  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-230 <GLA>  
A:Cross-references: UNIPROT:O92PT9; GB:AL592022; PIDN:CA95955.1; PID:g16413175; GSPDB:C  
A:Experimental source: strain C1p11262  
C:Genetics:  
A:Gene: lin0723

Query Match 26.9%; Score 69.5; DB 2; Length 230;  
Best Local Similarity 30.5%; Pred. No. 1.6;  
Matches 18; Conservative 10; Mismatches 16; Indels 15; Gaps 1;

Oy 2 EKETPVQK-----CORLQIEMNELINEVALQVDRKVADEEKOSYDAV 45  
Db 33 EVESYSEKLEQLLENHOKELERKMSAIEIQOKLANERAAALKAERQALIEIKOSAEAI 91

RESULT 3  
H72670  
hypothetical protein APE0790 - *Aeropyrum pernix* (strain K1)  
C:Species: *Aeropyrum pernix*  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: H72670  
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hatake, Y.; Jin-no, K.; Takah  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K  
awa, Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic *Crenarchaeon*, *Aeropyr*  
A:Reference number: A72450; MUID:99310339; PMID:10382966  
A:Accession: H72670  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-297 <KAN>  
A:Cross-references: UNIPROT:Q9YDX9; DDBJ:AP000060; NID:g5104188; PIDN:BAA79768.1; PID:d1

A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE0790

Query Match 26.4%; Score 68; DB 2; Length 297;  
Best Local Similarity 46.7%; Pred. No. 3.1;  
Matches 14; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Oy 12 RLQIEMNELINEVALQVDRKVADEEKOSY 41  
Db 29 RLKEERKLINIEVALKEERKASREKREY 58

RESULT 4  
H70314  
hypothetical protein aq\_157 - *Aquifex aeolicus*  
C:Species: *Aquifex aeolicus*  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
C:Accession: H70314  
R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov  
v.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.  
A:Reference number: A70300; MUID:98196666; PMID:9537320  
A:Accession: H70314  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-162 <AGP>  
A:Cross-references: UNIPROT:O66547; GB:AE000676; NID:g2982884; PIDN:AAC06512.1; PID:g298;  
A:Experimental source: strain VFS  
C:Genetics:  
A:Gene: aq\_157

Query Match 26.2%; Score 67.5; DB 2; Length 162;  
Best Local Similarity 35.3%; Pred. No. 1.9;  
Matches 18; Conservative 13; Mismatches 19; Indels 1; Gaps 1;

Oy 3 KETPVQKQCORLQIEMNELINEVALQVDRKVADEEKOSYDAVAVTVISTAR 53  
Db 81 KEKAKAIEQLQEDRLKQLQVE-AQSKSRKKAIEKRVKPVKIVIVESTAK 130

RESULT 5  
C84799  
hypothetical protein At2g37960 [imported] - *Arabidopsis thaliana*  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: C84799  
R:Jin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; J  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: C84799  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-508 <STO>  
A:Cross-references: UNIPROT:Q7XJQ8; GB:AE002093; NID:g4895188; PIDN:AA032775.1; GSPDB:GNI  
A:Genetics:  
A:Gene: At2g37960  
A:Map position: 2

Query Match 25.6%; Score 66; DB 2; Length 508;  
Best Local Similarity 34.2%; Pred. No. 9.1;  
Matches 13; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

Oy 5 TPVQKQCORLQIEMNELINEVALQVDRKVADEEKOSYD 42  
Db 55 SPAREVNSLLPLINEIINIEYIRLKKKIVMDQESKID 92

RESULT 6

T38346  
 hypothetical protein SPAC24C9.05c - fission yeast (Schizosaccharomyces pombe)  
 C/Species: Schizosaccharomyces pombe  
 C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 A/Accession: T38346  
 R/Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
 submitted to the EMBL Data Library, August 1995  
 A/Reference number: Z21787  
 A/Accession: T38346  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-730 <MDS>  
 A/Cross-references: UNIPROT:O13965; EMBL:Z98601; PDB:1CAB1262.1; GSPDB:GN00066; SPDB:SF  
 A/Experimental source: strain 972h-; cosmid c24C9  
 C/Genetics:  
 A/Gene: SPDB:SPAC24C9.05c  
 A/Map position: 1

Query Match 25.2%; Score 65; DB 2; Length 730;  
 Best Local Similarity 28.2%; Pred. No. 17;  
 Matches 20; Conservative 11; Mismatches 20; Indels 20; Gaps 2;

Qy 1 GKEKTPVQKCOR-----LQIEMNELNEVAALQVDRK-----VADEEKOS 40  
 Db 47 GNFPSPVRKIRNGEPGTVDNALDPAITVMOSLIVETATQMAKRONCVLVDDDEQL 106

Qy 41 YDAVAVATVIST 51  
 Db 107 AGIVATVDIAT 117

RESULT 7  
 B81688  
 conserved hypothetical protein TC0569 [imported] - Chlamydia muridarum (strain Nig9)  
 C/Species: Chlamydia muridarum, Chlamydia trachomatis Mopn  
 C/Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
 A/Accession: B81688  
 R/Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
 C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
 Nucleic Acids Res. 28, 1397-1406, 2000  
 A/Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.  
 A/Reference number: A81500; MUID:20150255; PMID:10684935  
 A/Accession: B81688  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-157 <TEF>  
 A/Cross-references: UNIPROT:Q9PK98; GB:AE002325; GB:AE002160; NID:g7190604; PDB:AAF3940  
 A/Experimental source: strain Nig9 (Mopn)  
 C/Genetics:  
 A/Gene: TC0569

Query Match 25.0%; Score 64.5; DB 2; Length 157;  
 Best Local Similarity 31.5%; Pred. No. 3.9;  
 Matches 17; Conservative 12; Mismatches 22; Indels 3; Gaps 2;

Qy 2 EKETPVQKCORLQIEMNELNEVAALQVDRKVADEEKOSYDAV--VATVISTAR 53  
 Db 24 EKIPVSPCKLEALQKFLKEVAVAK-DRYLQPRENVYSLPGVAIVKAR 76

RESULT 8  
 S00084  
 myosin heavy chain, fast skeletal muscle - rabbit (fragment)  
 C/Species: Oryctolagus cuniculus (domestic rabbit)  
 C/Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Jul-2004  
 A/Accession: S00084; S14807  
 R/Maeda, K.; Szekiel, G.; Wittinghofer, A.  
 Eur. J. Biochem. 167, 97-102, 1987  
 A/Title: Characterization of cDNA coding for the complete light meromyosin portion of a  
 A/Reference number: S00084; MUID:87304245; PMID:3305014  
 A/Accession: S00084  
 A/Molecule type: mRNA  
 A/Residues: 1-676 <MAE>

A/Cross-references: UNIPROT:P02562; EMBL:X05958; NID:g1622; PDB:1CAA29391.1; PID:g136424  
 A/Note: the sequence from Fig. 5 is inconsistent with that from Fig. 3 in having 561 Arg  
 R/Maeda, K.; Roesch, A.; Maeda, Y.; Kalbitzer, H.R.; Wittinghofer, A.  
 FEBS Lett. 281, 23-26, 1991  
 A/Title: Rabbit skeletal muscle myosin. Unfolded carboxyl-terminus and its role in molec  
 A/Reference number: S14807; MUID:91200294; PMID:2015900  
 A/Accession: S14807  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 668-676 <MA2>  
 C/Superfamily: myosin heavy chain; myosin motor domain homology  
 C/Keywords: actin binding; ATP; coiled coil; muscle; skeletal muscle  
 P:21-676/Domain: light meromyosin <LME>

Query Match 24.6%; Score 63.5; DB 2; Length 676;  
 Best Local Similarity 35.1%; Pred. No. 24;  
 Matches 20; Conservative 13; Mismatches 13; Indels 11; Gaps 3;

Qy 1 GKEKTPVQKCORLQIEMNELNEVAALQVDRKVA--DEE-----KQSYDAVAVATVIST 51  
 Db 290 GSLHBEKGIARIDLEINQVKS-----IDRKIAKDEIDQLKRNHLRVVDSMST 341

RESULT 9  
 S52696  
 myosin heavy chain - rainbow trout (fragment)  
 C/Species: Oncorhynchus mykiss (rainbow trout)  
 C/Date: 19-May-1995 #sequence\_revision 01-Sep-1995 #text\_change 09-Jul-2004  
 A/Accession: S52696  
 R/Gauvre, L.L.; Fauconneau, B.B.  
 submitted to the EMBL Data Library, March 1995  
 A/Description: Cloning of a fast skeletal myosin heavy chain expressed both in embryo an  
 A/Reference number: S52696  
 A/Accession: S52696  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-698 <GAV>  
 A/Cross-references: UNIPROT:Q91208; EMBL:Z48794; NID:g755770; PDB:1CAA88724.1; PID:g7557  
 A/Superfamily: myosin heavy chain; myosin motor domain homology  
 C/Keywords: ATP

Query Match 24.6%; Score 63.5; DB 2; Length 698;  
 Best Local Similarity 36.8%; Pred. No. 24;  
 Matches 21; Conservative 10; Mismatches 15; Indels 11; Gaps 3;

Qy 1 GKEKTPVQKCORLQIEMNELNEVAALQVDRKVA--DEE-----KQSYDAVAVATVIST 51  
 Db 312 GTLHBEKGIARIDLEINQVKS-----IDRKIAKDEIDQLKRNHLRVVDSMST 363

RESULT 10  
 JX0178  
 myosin heavy chain, fast skeletal muscle, adult [validated] - chicken  
 N/Contains: myosin ATPase (EC 3.6.4.1)  
 C/Species: Gallus gallus (chicken)  
 C/Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
 A/Accession: JX0050; JX0051; JX0178; A26365; S02082; P00009; S39081; S24351; S05  
 J/Hayashida, M.; Maica, T.; Matsuda, G.  
 J. Biochem. 110, 54-59, 1991  
 A/Title: The primary structure of skeletal muscle myosin heavy chain: I. Sequence of the  
 A/Reference number: JX0050; MUID:92041767; PMID:1939027  
 A/Accession: JX0050  
 A/Molecule type: protein  
 A/Residues: 1-205 <HAY>  
 A/Cross-references: UNIPROT:Q90913; UNIPROT:Q90909  
 R/Komine, Y.; Maica, T.; Matsuda, G.  
 J. Biochem. 110, 60-67, 1991  
 A/Title: The primary structure of skeletal muscle myosin heavy chain: II. Sequence of th  
 A/Reference number: JX0051; MUID:92041768; PMID:1939028  
 A/Accession: JX0051  
 A/Molecule type: protein  
 A/Residues: 206-636 <KOM>  
 R/Maica, T.; Miyashita, T.; Matsuzono, K.; Tanioaka, Y.; Matsuda, G.

J. Biochem. 110, 68-74, 1991  
A:Title: The primary structure of skeletal muscle myosin heavy chain: III. Sequence of t  
A:Reference number: EX0052; MUID:92041769; PMID:1935029  
A:Accession: EX0052  
A:Molecule type: protein  
A:Residues: 201-213;632-637 <MA1>  
J. Malta, T.; Yajima, E.; Nagata, S.; Miyaniishi, T.; Nakayama, S.; Matsuda, G.  
J. Biochem. 110, 75-87, 1991  
A:Title: The primary structure of skeletal muscle myosin heavy chain: IV. Sequence of th  
A:Reference number: UX0178; MUID:92041770; PMID:1935030  
A:Accession: UX0178  
A:Molecule type: protein  
A:Residues: 833-1938 <MA2>  
J. Malta, T.; Hayashida, M.; Tanioka, Y.; Komine, Y.; Matsuda, G.  
Proc. Natl. Acad. Sci. U.S.A. 84, 416-420, 1987  
A:Title: The primary structure of the myosin head.  
A:Reference number: A26365; MUID:87092420; PMID:3467365  
A:Accession: A26365  
A:Molecule type: protein  
A:Residues: 1-129, 'X', 131-139, 141-550, 'X', 552-754, 'X', 756-784, 'QL', 787-804, 806-810 <MA3>  
J. Matanabe, B.  
Biochem. Hoppe-Seyler 370, 55-61, 1989  
A:Title: Amino-acid sequence of the hinge region in chicken myosin subfragment-2.  
A:Reference number: S02082; MUID:89228549; PMID:2713098  
A:Accession: S02082  
A:Molecule type: protein  
A:Residues: 1144-1270 <MAT>  
R. Tajima, E.  
Nagasaki Igakkai Zasshi 65, 409-430, 1990  
A:Title: Study on tail region of skeletal muscle myosin; primary structure and protease  
A:Reference number: PM0009  
A:Accession: PM0009  
A:Molecule type: protein  
A:Residues: 1304-1938 <YAJ>  
R. Moore, L.A.; Arrizubieta, M.J.; Tidyman, W.E.; Herman, L.A.; Bandman, E.  
submitted to the EMBL Data Library, August 1991  
A:Description: Analysis of the chicken fast myosin heavy chain family: Localization of I  
A:Reference number: S39081  
A:Accession: S39081  
A:Molecule type: mRNA  
A:Residues: 1081-1203, 'DV', 1206-1342, 'E', 1344-1544, 'S', 1546-1795, 'HV', 1798-1829, 'S', 1831  
A:Cross-references: EMBL:M74084  
R. Moore, L.A.; Arrizubieta, M.J.; Tidyman, W.E.; Herman, L.A.; Bandman, E.  
J. Mol. Biol. 225, 1143-1151, 1992  
A:Title: Analysis of the chicken fast myosin heavy chain family. Localization of isoform  
A:Reference number: S24348; MUID:92309413; PMID:1377278  
A:Accession: S24351  
A:Molecule type: mRNA  
A:Residues: 1082-1182, 'T', 1184-1203, 'DV', 1206-1342, 'E', 1344-1544, 'S', 1546-1780, 'W', 1782-  
A:Cross-references: EMBL:M74084  
R. Matanabe, B.  
Biochem. Hoppe-Seyler 370, 1027-1034, 1989  
A:Title: Complete amino-acid sequence of subfragment-2 in adult chicken skeletal muscle  
A:Reference number: S05515; MUID:90121764; PMID:2610940  
A:Accession: S05515  
A:Molecule type: protein  
A:Residues: 842-906, 'Q', 908-1270 <MA3>  
R. Matanabe, B.  
Biochem. Hoppe-Seyler 370, 549-558, 1989  
A:Title: Amino-acid sequence of the short subfragment-2 in adult chicken skeletal muscle  
A:Reference number: S04501; MUID:89374803; PMID:2775482  
A:Accession: S04501  
A:Molecule type: protein  
A:Residues: 852-906, 'Q', 908-1108 <MA2>  
J. Matanabe, G.; Malta, T.; Miyaniishi, T.; Hayashida, M.  
J. Protein Chem. 6, 33-46, 1987  
A:Title: Structure and function of muscle myosin.  
A:Reference number: A60877  
A:Accession: A60877  
A:Molecule type: protein  
A:Residues: 1-139, 141-205 <MA1>  
R. Gulick, J.; Kropp, K.; Robbins, J.  
J. Biol. Chem. 260, 14513-14520, 1985

A:Title: The structure of two fast-white myosin heavy chain promoters. A comparative stu  
A:Reference number: A92507; MUID:86033956; PMID:2997212  
A:Accession: A24124  
A:Molecule type: DNA  
A:Residues: 'W', 1-168 <GUL>  
A:Cross-references: GB:M13512; GB:M12083; GB:M13510; NID:g212363; PIDN:AAA48966.1; PID:g  
J. Kropp, K.; Gulick, J.; Robbins, J.  
J. Biol. Chem. 261, 6613-6618, 1986  
A:Title: A canonical sequence organization at the 5'-end of the myosin heavy chain genes  
A:Reference number: A92587; MUID:86196091; PMID:3009465  
A:Accession: C25217  
A:Molecule type: DNA  
A:Residues: 'W', 1-56, 'T', 58-76, 'I', 78-168 <KRO>  
A:Cross-references: GB:M13515; GB:M13511; NID:g212373; PIDN:AAA48971.1; PID:g555468  
C:Comment: This is a fragment of the globular head.  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
C:Keywords: acetylated amino end; actin binding; ATP; coiled coil; hydrolase; methylated  
F:1-1938/Product: myosin heavy chain #status experimental <MAT>  
F:89-768/Domain: myosin motor domain homology <MMOT>  
F:179-186/Region: nucleotide-binding motif A (P-loop)  
F:550-587/Region: actin binding #status predicted  
F:657-679/Region: actin binding #status predicted  
F:841-1938/Domain: coiled coil <COI>  
F:841-1289/Region: S2  
F:852-1108/Domain: short subfragment 2 <SUB2>  
F:1290-1938/Region: light meromyosin  
F:1/Modified site: acetylated amino end (Ala) #status experimental  
F:35/Modified site: N6-methyllysine (Lys) #status experimental  
F:130,551/Modified site: N6,N6,N6-trimethyllysine (Lys) #status experimental  
F:185/Binding site: ATP (Lys) #status predicted  
F:697,707/Active site: Cys #status predicted  
F:755/Modified site: 3'-methylhistidine (His) #status experimental

Query Match 24.6%; Score 63.5; DB 1; Length 1938;  
Best Local Similarity 36.7%; Pred. No. 71;  
Matches 18; Conservative 13; Mismatches 7; Indels 11; Gaps 3;  
Oy 9 KCRLEFMNELLNEVALQVDRKVA--DEE---KQSYDAVATVIST 51  
Db 1559 KILRLQELNIOKSE-----IDRKIAKDERIDQLKKNHRIKIVSMOST 1602

RESULT 11  
F64596  
hypothetical protein HP0614 - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004  
C:Accession: F64596  
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKenney  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.  
Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.A.  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467; PMID:9252185  
A:Accession: F64596  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-111 <TOM>  
A:Cross-references: UNIPROT:O25335; GB:A8000576; GB:A8000511; NID:g2313736; PIDN:AA00768f  
C:Superfamily: Helicobacter pylori hypothetical protein jhp0557

Query Match 24.4%; Score 63; DB 2; Length 111;  
Best Local Similarity 40.6%; Pred. No. 4;  
Matches 13; Conservative 8; Mismatches 11; Indels 0; Gaps 0;  
Oy 11 ORLCIEMNELLNEVALQVDRKVADEKQSYD 42  
Db 12 QALQAKKKLITELAEFAIKVSSERKSSFN 43

RESULT 12  
E84283

C;Genet1CS:  
A:Gene: MHC

Qy	13	LQIEMELNEVALQVDRKVADEKOSYDAVV	45
	:::::	:::::	:::::
Db	59	IEIEQKLANEKALQAEQAIIEELRDAAEKEI	91
	:::::	:::::	:::::

24.0%; Score 62; DB 2; Length 230;  
 Best Local Similarity 36.4%; Pred. No. 11;  
 Matches 12; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: November 3, 2005, 21:37:29 ; Search time 93.246 Seconds  
(without alignments)  
291.060 Million cell updates/sec

Title: 09782816-52  
Perfect score: 258  
Sequence: 1 GKEKTPYQKQCRLOIEMNEL.....ADEKQSYDANVAIVISTAR 53

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	243.5	94.4	380	Q9V4Y9	Q9V4Y9 drosophila
2	127.5	49.4	388	Q7PZ25	Q7PZ25 anopheles g
3	82	31.8	402	Q9PTG6	Q9PTG6 gallus gall
4	77	29.8	242	Q6CYK1	Q6CYK1 erwinia car
5	75	29.1	405	Q7T3H1	Q7T3H1 brachydanio
6	74	28.7	401	DCRT2	DCRT2 mouse
7	74	28.7	402	Q6A7H5	Q6A7H5 mus musculu
8	73.5	28.5	403	Q66U30	Q66U30 xenopus lae
9	72.5	28.1	400	DCRT2	DCRT2 human
10	70.5	27.3	403	Q6IRB3	Q6IRB3 homo sapien
11	70.5	27.3	755	Q9JNMI	Q9JNMI arabidopsis
12	70.5	27.3	879	Q9YHDE	Q9YHDE rana catesb
13	70.5	27.3	1937	MYH8	MYH8 human
14	69.5	26.9	230	Q9ZDT9	Q9ZDT9 homo sapien
15	68	26.4	297	Q9YDX9	Q9YDX9 aeropyrum p
16	67.5	26.2	162	Y157	Y157 aquifex aeo
17	67.5	26.2	338	Q7XVY2	Q7XVY2 xenopus lae
18	67.5	26.2	1939	Q6DF06	Q6DF06 xenopus tro
19	66	25.6	480	Q8H1N5	Q8H1N5 arabidopsis
20	66	25.6	508	Q7XUQ8	Q7XUQ8 arabidopsis
21	65.5	25.4	1935	Q6NXX3	Q6NXX3 xenopus tro
22	65	25.2	730	YE45	YE45 schizosacch
23	64.5	25.0	157	Q9PK98	Q9PK98 chlamydia m
24	64.5	25.0	935	Q8I734	Q8I734 trypanosoma
25	64	24.8	885	RA50	RA50 sulfobus
26	63.5	24.6	657	Q8IS89	Q8IS89 spiroplasma
27	63.5	24.6	698	Q9I208	Q9I208 oncorhynch
28	63.5	24.6	708	Q9YHD7	Q9YHD7 rana catesb
29	63.5	24.6	826	Q9YHD5	Q9YHD5 rana catesb
30	63.5	24.6	1084	MYSS	MYSS rabbit
31	63.5	24.6	1937	Q8UIP5	Q8UIP5 oncorhynch

32	63.5	24.6	1938	1	MYSS CHICK	P13538 gallus gall
33	63.5	24.6	1941	1	MYH2_HUMAN	Q9UKX2 homo sapien
34	63.5	24.6	1941	2	Q86T56	Q86T56 homo sapien
35	63.5	24.6	2664	2	Q7RELO	Q7RELO plasmodium
36	63	24.4	111	2	Q25335	Q25335 helicobacter
37	63	24.4	207	2	Q983X1	Q983X1 rhizobium l
38	63	24.4	716	2	Q46317	Q46317 carnobacter
39	63	24.4	716	2	Q9REY3	Q9REY3 carnobacter
40	62.5	24.2	45	2	Q71I08	Q71I08 lactobacill
41	62.5	24.2	222	2	Q9H083	Q9H083 halobacteri
42	62.5	24.2	469	2	Q9ESM6	Q9ESM6 mus musculu
43	62.5	24.2	892	2	Q63939	Q63939 rattus sp.
44	62.5	24.2	1119	2	P87344	P87344 theagra ch
45	62.5	24.2	1725	2	Q7RT66	Q7RT66 plasmodium

## ALIGNMENTS

RESULT 1	ID	Q9V4Y9	PRELIMINARY;	PRT;	380 AA.
AC	Q9V4Y9				
DT	01-MAY-2000	(TREMBLrel. 13, Created)			
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)			
DT	25-OCT-2004	(TREMBLrel. 28, Last annotation update)			
DE	CG8269-PA (LD07994P)				
GN	Name=Dmu; ORFNames=CG8269;				
OS	Drosophila melanogaster (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephydroidea; Drosophilidae; Drosophila.				
NCBI	TaxID=7227;				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;				
RA	Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,				
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA	Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,				
RA	Branden R.C., Rogers Y.H., Blazey R.G., Champe M., Pfeiffer B.D.,				
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,				
RA	Abtill J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,				
RA	Bellaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,				
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,				
RA	Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,				
RA	Buttis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,				
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,				
RA	de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,				
RA	Dodson K., Dou P.L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,				
RA	Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,				
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,				
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,				
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,				
RA	Hoskins D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,				
RA	Jatelli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,				
RA	Jatelli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,				
RA	Laeso P., Lei Y., Levitsky A.A., Li Z., Li Z., Liang Y., Lin X.,				
RA	Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,				
RA	Mekullov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,				
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,				
RA	Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Paclet J.M.,				
RA	Palazzo K.M., Pittman G.S., Pan S., Pollard J.F., Puti V., Reese M.G.,				
RA	Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,				
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,				
RA	Spier B., Spradling A.C., Stapleton M., Strong R., Sun E.,				
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,				
RA	Wang Z.Y., Wassarman D.A., Weisscock G.M., Weissbach J.,				
RA	Williams S.M., Woodgerger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,				
RA	Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,				
RA	Zhang X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,				
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,				
RT	"The genome sequence of Drosophila melanogaster."				

```
RL Science 287:2185-2195 (2000) .
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Ceiniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champagne M., Dugan S.P., Friese E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svitskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weissstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
RT "finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002) .
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svitskas R.,
RA Patel S., Friese E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Ceiniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomic perspective."
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002) .
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Miura S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hrdceky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Betencourt B.R., Ceiniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yanada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review."
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002) .
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Friese E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Ceiniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003835; AAF59034.1; -.
DR EMBL; AY061092; AAL28640.1; -.
DR IncAct; Q9V4Y9; -.
DR FlyBase; FBgn0021825; Dm.
DR GO; GO:0005869; C:dynactin complex; IEA.
DR GO; GO:0007017; P:microtubule-based process; IEA.
DR InterPro; IPR006996; DnaMittin.
DR Pfam; PF04912; DnaMittin; 1.
SQ SEQUENCE 380 AA; 41998 MW; CF7E1D3BFF5989C5 CRC64;

Query Match 94.4%; Score 243.5; DB 2; Length 380;
Best Local Similarity 98.1%; Pred. No. 16-18; Mismatches 0; Indels 1; Gaps 1;
Matches 52; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 GEEKTPVOKCORLQIEMNELINEVALQVDRKVADEKQSYDAVATVISTAR 53
Db 94 GEEKTPVOKCORLQIEMNELINEVALQVDRKVADEKQSYDAVATVISTAR 145

RESULT 2
Q7P225.
```

```
ID Q7P225 PRELIMINARY; PRT; 388 AA.
AC Q7P225;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE AGCP9561 (Fragment).
GN Name=agCG52050; ORFNames=ENSG00000015420;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB0100886; EAA00075.1; -.
DR GO; GO:0005869; C:dynactin complex; IEA.
DR GO; GO:0007017; P:microtubule-based process; IEA.
DR InterPro; IPR006996; DnaMittin.
DR Pfam; PF04912; DnaMittin; 1.
FT NON TER 1
FT SEQUENCE 388 AA; 43021 MW; 4F5A549B51D52555 CRC64;

Query Match 49.4%; Score 127.5; DB 2; Length 388;
Best Local Similarity 49.1%; Pred. No. 7-5e-06;
Matches 26; Conservative 11; Mismatches 15; Indels 1; Gaps 1;

Qy 1 GEEKTPVOKCORLQIEMNELINEVALQVDRKVADEKQSYDAVATVISTAR 53
Db 94 GEEKTPVOKCORLQIEMNELINEVALQVDRKVADEKQSYDAVATVISTAR 145

RESULT 3
Q9PTG6 PRELIMINARY; PRT; 402 AA.
ID Q9PTG6;
AC Q9PTG6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE DnaMittin.
GN Name=p50;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20056093; PubMed=10588646;
RA Valetti C., Wetzel D.M., Schrader M., Hasbani M.J., Gill S.R.,
RA Kreis T.E., Schroer T.A.;
RT "Role of dynactin in endocytic traffic: effects of dynactin
RT overexpression and colocalization with CLIP-170."
RL Mol. Biol. Cell 10:4107-4120(1999) .
RN [2]
RP SEQUENCE FROM N.A.
RX Schroer T.A., Gill S.R., Hasbani J., Grego C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF200744; AAT13996.1; -.
DR GO; GO:0005869; C:dynactin complex; IEA.
DR GO; GO:0007017; P:microtubule-based process; IEA.
DR InterPro; IPR006996; DnaMittin.
DR Pfam; PF04912; DnaMittin; 1.
SQ SEQUENCE 402 AA; 45126 MW; F229C467C630DCB9 CRC64;

Query Match 31.8%; Score 82; DB 2; Length 402;
Best Local Similarity 48.7%; Pred. No. 0.86;
Matches 19; Conservative 6; Mismatches 12; Indels 2; Gaps 1;
```



```

Qy 1 GEEKTPVQKCORLQIENNELNEVVALQVDRKVADEKQ 37
Db 96 GAKETPOQRYORLQHEVQELIRDEQIQSAVESASAE 134

RESULT 4
ID 06CYK1 PRELIMINARY; PRT; 242 AA.
AC 06CYK1;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=BCA4506;
OS Erwinia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
CX NCBI_TaxID=29471;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCRI 1043 / ATCC BAA-672;
RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebahia M., Pritchard L., Holden M.T.G., Hyman L.J.,
RA Holdev M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
RA Atkin R., Bacon N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagsels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
RT "Genome sequence of the enterobacterial phytopathogen Erwinia
RT carotovora subsp. atroseptica and characterization of virulence
RT factors."
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
RM EMBL; BX350851; CAG77401.1;
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 242 AA; 26438 MW; A67C186B12B6240C CRC64;

Query Match 29.8%; Score 77; DB 2; Length 242;
Best Local Similarity 40.9%; Pred. No. 1.8;
Matches 18; Conservative 8; Mismatches 18; Indels 0; Gaps 0;

Qy 4 ETPVQKCORLQIENNELNEVVALQVDRKVADEKQSDAVAT 47
Db 36 EQPDKCRHLPMMARRRINQSLAVDSGLALHRQSDALVFT 79

RESULT 5
ID 07T3H1 PRELIMINARY; PRT; 405 AA.
AC 07T3H1;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Similar to dynactin 2 (P50).
GN ORFNames=zgc:63867;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Danio.
CX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshitsuki S., Caramini P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

```

```

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek A.J., Small D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Strausberg R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC053120; AAH53120.1;
DR ZFIN; ZDB-GENE-040426-1279; zgc:63867.
DR GO; GO:0005669; C:dynactin complex; IEA.
DR GO; GO:007017; P:microtubule-based process; IEA.
DR InterPro; IPR06996; Dynactin.
DR Pfam; PF04912; Dynactin; 1.
SQ SEQUENCE 405 AA; 44625 MW; CA00047342500953 CRC64;

Query Match 29.1%; Score 75; DB 2; Length 405;
Best Local Similarity 43.6%; Pred. No. 5.2;
Matches 17; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GEEKTPVQKCORLQIENNELNEVVALQVDRKVADEKQ 39
Db 96 GAKETPOQRYORLQHEVQELIRDEQIQSAVESASAE 134

RESULT 6
ID DCT2 MOUSE STANDARD; PRT; 401 AA.
AC 09PKJ8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Dynactin complex 50 kDa subunit (50 kDa dynein-associated polypeptide)
DE (p50 dynactin) (DCTN-50) (Dynactin 2) (Growth cone membrane protein
DE 23-48k) (GMP23-48k).
GN Name=DCTN2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshitsuki S., Caramini P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek A.J., Small D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE OF 65-74; 77-91; 102-116; 156-170; 194-216 AND 309-320,
RP SUBCELLULAR LOCATION, AND DEVELOPMENTAL STAGE.

```

```

CC TISSUE=Brain;
RX MEDLINE=97289622; PubMed=9144527; DOI=10.1006/dbrc.1997.6447;
RA Abe T.K., Tanaka H., Iwanaga T., Odani S., Kuwano R.;
RT "The presence of the 50-kDa subunit of dynactin complex in the nerve
RT growth cone.";
RL Biochem. Biophys. Res. Commun. 233:295-299(1997).
RN [3]
RP INTERACTION WITH BICD2.
RX MEDLINE=21376052; PubMed=11483508; DOI=10.1093/emboj/20.15.4041;
RA Hoogenraad C.C., Akhmanova A., Howell S.A., Dottiand B.R.,
RA de Zeeuw C.I., Willemsen R., Visser P., Grosveld F., Galjart N.;
RT "Mammalian Golgi-associated Bicaudal-D2 functions in the dynein-
RT dynein pathway by interacting with these complexes.";
RL EMO J. 20:4041-4054(2001).
CC -1- FUNCTION: Modulates cytoplasmic dynein binding to an organelle,
CC and plays a role in prometaphase chromosome alignment and spindle
CC organization during mitosis. May play a role in karyopse formation
CC during brain development.
CC -1- SUBUNIT: Subunit of dynactin, a multiprotein complex associated
CC with dynein (By similarity). Interacts with BICD2.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
CC -1- DEVELOPMENTAL STAGE: Present at high levels in both cytoplasmic
CC and membrane-associated forms in neonates. Levels of membrane-
CC associated form are greatly reduced in the adult.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: BC004613; AAH04613.1; -
DR MGD: MGI:107733; Dctn2.
DR InterPro: IPR006996; Dynameitin.
DR Pfam: PF04912; Dynameitin; 1.
KW Coiled coil; Cytoskeleton; Direct protein sequencing; Dynein;
KW Membrane; Microtubule; Motor protein.
FT INT MET 0 By similarity.
FT DOMAIN 98 131 Coiled coil (Potential).
FT DOMAIN 214 244 Coiled coil (Potential).
SQ SEQUENCE 401 AA; 43985 MW; 1535E4ABD5940EBC CRC64;

Query Match 28.7%; Score 74; DB 1; Length 401;
Best Local Similarity 35.8%; Pred. No. 6.6; Indels 0; Gaps 0;
Matches 19; Conservative 10; Mismatches 24;

Qy 1 GKEETPVQKQRLQIEMNELNEVAALQVDRKVADEKQSYDAVAVTISTAR 53
Db 93 GVKETPQOKYQRLHVEVQELTTEVEKIKTYKESATSEKLPVVLAKQLALAK 145

RESULT 7
ID 06AYH5 PRELIMINARY; PRT; 402 AA.
AC 06AYH5;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Dynein 2.
GN Name=Dctn2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shmamen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

```

```

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Ditschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stepietson M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Uedini T.B., Toshiyuki S., Carrinini P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Heitlen E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shervenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M., Buterfield Y.S.,
RA Krywinski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RC Director MGC Project;
RA Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC079042; AAH79042.1; -
DR GO: GO:0005869; C:dynactin complex; IEA.
DR GO: GO:0007017; P:microtubule-based process; IEA.
DR InterPro: IPR006996; Dynameitin; 1.
DR Pfam: PF04912; Dynameitin; 1.
SQ SEQUENCE 402 AA; 44148 MW; 55033535A4FB052 CRC64;

Query Match 28.7%; Score 74; DB 2; Length 402;
Best Local Similarity 35.8%; Pred. No. 6.7;
Matches 19; Conservative 10; Mismatches 24; Indels 0; Gaps 0;

Qy 1 GKEETPVQKQRLQIEMNELNEVAALQVDRKVADEKQSYDAVAVTISTAR 53
Db 94 GVKETPQOKYQRLHVEVQELTTEVEKIKTYKESATSEKLPVVLAKQLALAK 146

RESULT 8
ID 06GJ30 PRELIMINARY; PRT; 403 AA.
AC 06GJ30;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE MGC82128 protein.
GN Name=MGC82128;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shmamen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Ditschenko L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedini T.B., Toshiyuki S., Carrinini P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

```

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Raley U., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalhus D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Klein S., Gerthard D.S.,  
 RL Submitted (Aug-2004) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; BC081081; AAH81081.1; -  
 DR InterPro; IPR006996; Dynameitin.  
 DR Pfam; PF04912; Dynameitin; 1.  
 SQ SEQUENCE 403 AA; 44737 MW; 86BE8CE54325EF3 CRC64;  
 Query Match 28.5%; Score 73.5; DB 2; Length 403;  
 Best Local Similarity 45.1%; Pred. No. 7.6;  
 Matches 23; Conservative 5; Mismatches 22; Indels 1; Gaps 1;  
 Oy 1 GSEKTPQKCORLQIENNELNEVAALQ-VDRKVADEKSYDAVAVTIS 50  
 Db 94 GKMETPOOKYORLHVEQLTQVEKTSWAKSEKLPVLAQKVAS 144  
 RESULT 9  
 DCT2 HUMAN STANDARD; PRT; 400 AA.  
 ID OI3561; O96VY2; Q9BW17;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 05-JUL-2004 (Rel. 44, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Dynameitin complex 50 kDa subunit (50 kDa dynein-associated polypeptide)  
 DE (50 dynameitin) (DCTN-50) (Dynameitin 2).  
 GN Name=DCTN2; Synonyms=DCTN50;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_taxonomy=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=96178072; PubMed=6647893; DOI=10.1083/jcb.132.4.617;  
 RA Echeverri C.J., Paschal B.M., Vaughan K.T., Vallee R.B.;  
 RT "Molecular characterization of the 50-kD subunit of dynein reveals  
 RT organization for the complex in chromosome alignment and spindle  
 RT organization during mitosis."  
 RL J. Cell Biol. 132:617-633(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta, Skin, and Uterus;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stepien M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carinci P., Prange C.,  
 RA Rana S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Raley U., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,

RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RP SEQUENCE OF 1-384 FROM N.A.  
 RA Aunais J.P., Yu-Jee L.-Y.;  
 RT "Human 50 kD dynein subunit, p50 dyneitin, isolated from HeLa  
 RT cells."  
 RL Submitted (NOV-2002) to the EMBL/Genbank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 1-13.  
 RC TISSUE=Platelet;  
 RX MEDLINE=22608298; PubMed=12665801; DOI=10.1038/nbt810;  
 RA Gevaert K., Goethals M., Martens L., Van Damme J., Staes A.,  
 RA Thomas G.R., Vandeckerkhove J.;  
 RT "Exploring proteomes and analyzing protein processing by mass  
 RT spectrometric identification of sorted N-terminal peptides."  
 RL Nat. Biotechnol. 21:566-569(2003).  
 CC -1- FUNCTION: Modulates cytoplasmic dynein binding to an organelle,  
 CC and plays a role in prometaphase chromosome alignment and spindle  
 CC organization during mitosis. May play a role in synapse formation  
 CC during brain development.  
 CC -1- SUBUNIT: Subunit of dynein, a multiprotein complex associated  
 CC with dynein.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; U50733; AAC50423.1; -  
 DR EMBL; BC000718; AAH00718.1; -  
 DR EMBL; BC000948; AAH09468.1; -  
 DR EMBL; BC014083; AAH14083.1; -  
 DR EMBL; AY189155; AAC34395.1; -  
 DR Genem; HGNC:2712; DCTN2.  
 DR MIM; 607376; -  
 DR GO; GO:0005813; C:centrosome; TAS.  
 DR GO; GO:0005869; C:dynein complex; TAS.  
 DR GO; GO:0000776; C:kinesin complex; TAS.  
 DR GO; GO:0008283; P:cell proliferation; TAS.  
 DR GO; GO:0007067; P:mitosis; TAS.  
 DR InterPro; IPR006996; Dynameitin.  
 DR Pfam; PF04912; Dynameitin; 1.  
 KW Coiled coil; Cytoskeleton; Direct protein sequencing; Dynein;  
 KW Membrane; Microtubule; Motor protein.  
 FT INIT MET 0  
 FT DOMAIN 98 131 Coiled coil (potential).  
 FT DOMAIN 213 243 Coiled coil (potential).  
 FT DOMAIN 378 398 Coiled coil (potential).  
 FT CONFLICT 34 34 A -> AFAOEL (in Ref. 1).  
 FT CONFLICT 35 35 E -> BLE (in Ref. 3).  
 FT CONFLICT 381 384 LATV -> PGHS (in Ref. 3).  
 SQ SEQUENCE 400 AA; 44099 MW; 0A95AE950CB270F CRC64;  
 Query Match 28.1%; Score 72.5; DB 1; Length 400;  
 Best Local Similarity 48.7%; Pred. No. 9.7;  
 Matches 19; Conservative 5; Mismatches 14; Indels 1; Gaps 1;  
 Oy 1 GSEKTPQKCORLQIENNELNEVAALQVD-RKVADEK 38  
 Db 93 GKMETPOOKYORLHVEQLTQVEKTSWAKSEKLPVLAQKVAS 131  
 RESULT 10  
 ID O6IRB3 PRELIMINARY; PRT; 403 AA.  
 AC O6IRB3;  
 AC O6IRB3;

DT 05-JUL-2004 (TReMBLrel. 27, Created)  
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
DE Dctn2-prov protein.  
GN Name=dctn2-prov.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
OC Xenopodidae; Xenopus.  
NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L.H., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheefter C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carrinci P., Prange C.,  
RA Raba S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McKernan P.U., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards D.K., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley A.C., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smalke D.B., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.,  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT Initiative.";  
RL Dev. Dyn. 225:384-391 (2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RA Klein S., Strausberg R.,  
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC070987; AAH70987.1; -  
DR GO; GO:0005869; C:dynactin complex; IEA.  
DR GO; GO:0007017; P:microtubule-based process; IEA.  
DR InterPro; IPR006996; Dynactin.  
DR Pfam; PF04912; Dynactin; 1.  
SQ SEQUENCE 403 AA; 44828 MW; 9FD158C7C7983062 CRC64;  
Query Match 27.3%; Score 70.5; DB 2; Length 403;  
Best Local Similarity 51.3%; Pred. No. 16;  
Matches 20; Conservative 3; Mismatches 15; Indels 1; Gaps 1;  
QY 1 GKKETPVQKQRLQIEMNELINEVAALQVD-RKYADEEK 38  
Db 94 GIKETPOOKYQRLHVEQLTQVEVKTKSTLKSATK 132  
RESULT 11  
Q9LNM1 PRELIMINARY; PRT; 755 AA.  
AC Q9LNM1;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
DE F12K21.7.  
OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,  
RA Kim C., Altairi H., Bei B., Chin C., Chlou J., Choi E.,  
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,  
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharbeky N.,  
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
RA Thavert A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
RA Ecker J.R.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
RA Khan S., Kim C., Altairi H., Bei B., Chin C., Chlou J., Choi E.,  
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,  
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharbeky N.,  
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
RA Thavert A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
RA Theologis A., Ecker J.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC023279; AAF9274.1; -  
DR InterPro; IPR001005; Myb DNA binding.  
DR PROSITE; PS00037; MYB 1; UNKNOWN 1.  
SQ SEQUENCE 755 AA; 83720 MW; B4DBCB410208022A CRC64;  
Query Match 27.3%; Score 70.5; DB 2; Length 755;  
Best Local Similarity 36.5%; Pred. No. 31;  
Matches 20; Conservative 9; Mismatches 16; Indels 7; Gaps 2;  
QY 2 EKETPVQKQRLQIEMNELINEVAALQVD-RKYADEEKQSYAVATYSTAR 53  
Db 582 EASRIQAAER---EKNEALSEAAAKLER---EEGRGRSEVAAAIQTTR 626  
RESULT 12  
Q9YHD8 PRELIMINARY; PRT; 879 AA.  
AC Q9YHD8;  
DT 01-MAY-1999 (TReMBLrel. 10, Created)  
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
DE Myosin heavy chain (Fragment).  
GN Name=MHC-1;  
OS Rana catesbeiana (Bull frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.  
NCBI\_TaxID=8400;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Tail muscle;  
RX MEDLINE=99179235; PubMed=10079518;  
RA DOI=10.1002/(SICI)1520-6408(1999)24:1/2<151::AID-DVGA4>3.3.CO;2-W;  
RA Hu H., Merrifield P., Atkinson B.G.;  
RT "Expression of the myosin heavy chain genes in the tail muscle of  
RT thyroid hormone-induced metamorphosing Rana catesbeiana tadpoles.";  
RL Dev. Genet. 24:151-164 (1999).  
DR EMBL; AF097904; AAD13769.1; -  
DR HSSP; P25054; IDEB.  
DR GO; GO:0016459; C:myosin; IEA.  
DR GO; GO:0003774; F:motor activity; IEA.  
DR InterPro; IPR002928; Myosin tail.  
DR InterPro; IPR000533; Tropomyosin.  
DR Pfam; PF01576; Myosin tail 1; 1.  
DR PRINTS; PR00194; TROPOMYOSIN.  
FT NON TER 1  
SQ SEQUENCE 879 AA; 101710 MW; 1C456851E968A3D8 CRC64;  
Query Match 27.3%; Score 70.5; DB 2; Length 879;  
Best Local Similarity 36.8%; Pred. No. 36;

Matches 21; Conservative 11; Mismatches 14; Indels 11; Gaps 3;

QY 1 GSEKTPVQKCRLOIENNELNEVALLOVDRKVA--DEE---KQSYAVVATVIST 51  
 Db 489 GSLHEESKIRIOLIELNOLKSE-----VDRKIAEKDEIEIQLKKNQRIDTQST 540

RESULT 13

MYH8 HUMAN STANDARD; PRT, 1937 AA.

AC P13535; Q14910;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 25-JAN-2005 (Rel. 46, Last annotation update)  
 DE Myosin heavy chain, skeletal muscle, perinatal (MYHC-perinatal).  
 GN Name=MYH8;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=90323631; PubMed=2373371; DOI=10.1016/0378-1119(90)90020-R;  
 RA Karsch-Mizrachi I., Peghali R., Shows T.B. Jr., Leitman L.A.;  
 RT "Generation of a full-length human perinatal myosin heavy-chain-  
 encoding cDNA.";  
 RL Gene 89:289-294(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=95324556; PubMed=7601129;  
 RA Ullian E.H., Kelly A.M., Pompidou A.J., Hoffman R., Schiaffino S.,  
 RA Steadman H.H., Rubinstein N.A.;  
 RT "Characterization of a human perinatal myosin heavy-chain  
 transcript.";  
 RL Eur. J. Biochem. 230:1001-1006(1995).  
 RN [3]  
 RP SEQUENCE OF 502-1937 FROM N.A.  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=90235862; PubMed=1691980;  
 RA Bober E., Buchberger-Seidl A., Braun T., Singh S., Goede H.W.,  
 RA Arnold H.H.;  
 RT "Identification of three developmentally controlled isoforms of human  
 myosin heavy chains.";  
 RL Eur. J. Biochem. 189:55-65(1990).  
 RN [4]  
 RP SEQUENCE OF 860-1937 FROM N.A.  
 RX MEDLINE=89234168; PubMed=2715179; DOI=10.1083/jcb.108.5.1791;  
 RA Peghali R., Leitman L.A.;  
 RT "Molecular genetic characterization of a developmentally regulated  
 human perinatal myosin heavy chain.";  
 RL J. Cell Biol. 108:1791-1797(1989).  
 RN [5]  
 RP SEQUENCE OF 1-46 FROM N.A.  
 RA Esser K., Tichar A., Myszkowski M.;  
 RT "Isolation and characterization of the human perinatal MHC promoter.";  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP VARIANT CARNEY COMPLEX VARIANT/TRISOMY-PSEUDOCAMPDODACTYL SYNDROME  
 RP GUN-674.  
 RX PubMed=15282353; DOI=10.1056/NEJMoA040584;  
 RA Veugelers M., Bressan M., McDermott D.A., Weremowicz S., Morton C.C.,  
 RA Mabry C.C., Lefevre J.-F., Zunnan A., Deestre A., Chaudron J.-M.,  
 RA Baeson C.T.;  
 RT "Mutation of perinatal myosin heavy chain associated with a Carney  
 RT complex variant.";  
 RL N. Engl. J. Med. 351:460-469(2004).  
 CC -1- FUNCTION: Muscle contraction.  
 CC -1- SUBUNIT: Muscle myosin is a hexameric protein that consists of 2  
 CC heavy chain subunits (MHC), 2 alkali light chain subunits (MLC)  
 CC and 2 regulatory light chain subunits (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.

CC -1- DOMAIN: The rodlike tail sequence is highly repetitive, showing  
 CC cycles of a 28-residue repeat pattern composed of 4 heptapeptides,  
 CC characteristic for alpha-helical coiled coils.  
 CC -1- DISEASE: Defects in MYH8 are a cause of Carney complex variant  
 CC (MIM:60837). Carney complex is a multiple neoplasia syndrome  
 CC characterized by spotty skin pigmentation, cardiac and other  
 CC myxomas, endocrine tumors, and psammomatous melanotic schwannomas.  
 CC Familial cardiac myxomas are associated with spotty pigmentation  
 CC of the skin and other phenotypes, including primary pigmented  
 CC nodular adrenocortical dysplasia, extracardiac (frequently  
 CC cutaneous) myxomas, schwannomas, and pituitary, thyroid,  
 CC testicular, bone, ovarian, and breast tumors. Cardiac myxomas do  
 CC not develop in all patients with the Carney complex, but affected  
 CC patients have at least two features of the complex or one feature  
 CC and a clinically significant family history.  
 CC -1- DISEASE: Defects in MYH8 are a cause of trismus-  
 CC pseudocampodactyly syndrome (MIM:18300); also called Hecht-Beals  
 CC or Dutch-Kentucky syndrome. The trismus-pseudocampodactyly  
 CC syndrome is a hereditary distal arthroproposis characterized by an  
 CC inability to open the mouth fully (trismus) and  
 CC pseudocampodactyly in which wrist dorsiflexion, but not  
 CC volarflexion, produces involuntary flexion contracture of distal  
 CC and proximal interphalangeal joints. Such hand and jaw  
 CC contractures are caused by shortened flexor muscle-tendon units.  
 CC Similar lower-limb contractures also produce foot deformity. The  
 CC trismus-pseudocampodactyly syndrome is a morbid autosomal  
 CC dominant trait with variable expressivity but high penetrance. In  
 CC these patients, trismus complicates dental care, feeding during  
 CC infancy, and intubation for anesthesia, and the  
 CC pseudocampodactyly impairs manual dexterity, with consequent  
 CC occupational and social disability. Many patients require surgical  
 CC correction of contractures.  
 CC -1- MISCELLANEOUS: Each myosin heavy chain can be split into 1 light  
 CC meromyosin (LMW) and 1 heavy meromyosin (HWM). It can later be  
 CC split further into 2 globular subfragments (S1) and 1 rod-shaped  
 CC subfragment (S2).  
 CC -1- SIMILARITY: Contains 1 IQ domain.  
 CC -1- SIMILARITY: Contains 1 myosin-like globular head domain.  
 CC -1- SIMILARITY: Contains 1 myosin-like globular head domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL; M36769; AAC17185.1; -;  
 CC EMBL; Z38133; CA86293.1; -;  
 CC EMBL; X51592; CAA35941.1; -;  
 CC EMBL; AF067143; AAC21557.1; -;  
 CC PIR; I38055; I38055.  
 CC HSSP; P13538; 2MYS.  
 CC Genew; HGNC:7578; MYH8.  
 CC MIM; 160741; -;  
 CC MIM; 60837; -;  
 CC MIM; 158300; -;  
 CC GO; GO:0005859; C:muscle myosin; TAS.  
 CC GO; GO:0008307; F:structural constituent of muscle; TAS.  
 CC InterPro; IPR000048; IQ\_region.  
 CC InterPro; IPR001609; Myosin\_head.  
 CC InterPro; IPR004009; Myosin\_N.  
 CC InterPro; IPR002928; Myosin\_tail.  
 CC Pfam; PF00612; IQ; 1.  
 CC Pfam; PF00063; Myosin\_head; 1.  
 CC Pfam; PF02736; Myosin\_N; 1.  
 CC Pfam; PF01576; Myosin\_tail; 1.  
 CC PRINTS; PR00193; MYOSINHEAVY.  
 CC PRODOM; PD000355; Myosin\_head; 1.  
 CC PROSITE; PS50096; IQ; 1.  
 CC Actin-binding; ATP-binding; Calmodulin-binding; Coiled coil;  
 CC Disease mutation; Methylation; Multigene family; Muscle protein;  
 CC Myosin; Thick filament.

FT DOMAIN 1 780 Myosin head-like.  
FT DOMAIN 781 813 IO.  
FT DOMAIN 842 1937 Coiled coil (potential).  
FT NP\_BIND 181 188 ATP.  
FT DOMAIN 658 680 Actin-binding.  
FT DOMAIN 760 774 Actin-binding.  
FT MOD\_RES 132 132 N6,N6,N6-trimethyllysine (potential).  
FT VARIANT 674 674 R -> Q (in Carney complex variant and trismus-pseudocamptodactyly syndrome).  
/FTID=VAR\_019810  
FT CONFLICT 15 15 A -> R (in Ref. 2).  
FT CONFLICT 970 970 E -> Q (in Ref. 1 and 4).  
FT CONFLICT 1072 1072 M -> N (in Ref. 3).  
FT CONFLICT 1247 1247 N -> H (in Ref. 1 and 4).  
FT CONFLICT 1251 1252 MC -> DGG (in Ref. 3).  
FT CONFLICT 1261 1261 E -> G (in Ref. 1 and 4).  
FT CONFLICT 1297 1297 K -> Q (in Ref. 1 and 4).  
FT CONFLICT 1377 1378 KY -> NT (in Ref. 3).  
FT CONFLICT 1504 1505 EN -> AH (in Ref. 1 and 4).  
FT CONFLICT 1847 1847 E -> D (in Ref. 1 and 4).  
FT CONFLICT 1914 1914 D -> H (in Ref. 2).  
SQ SEQUENCE 1937 AA; 222762 MW; A3ER2D151792B9B8 CRC64;

Query Match 27.3%; Score 70.5; DB 1; Length 1937;  
Best Local Similarity 28.9%; Pred. No. 80;  
Matches 22; Conservative 15; Mismatches 8; Indels 31; Gaps 4;

Qy 2 EKERPVKQ-----RLQIEMNELLNEVALQVDRKVA--DEE-- 37  
Db 1533 KKQVEQCEIQALIEAEASLEHGEKILRIQLLNQKSE---VDRKIAEKDEID 1587  
Qy 38 --KQSYDAVAVATVIST 51  
Db 1588 QLKXNHRVETMOST 1603

RESULT 14  
ID Q92DT9 PRELIMINARY; PRT; 230 AA.  
AC Q92DT9;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Lin0723 protein.  
GN OrderedLocustNames=lin0723;  
OS Listeria innocua.  
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
OX NCBI\_TaxID=1642;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CLIP 11262 / Serovar 6a;  
RX MEDLINE=21537279; PubMed=1093669; DOI=10.1126/science.1063447;  
RA Glaser P., Frangoul L., Buchrieser C., Rusniok C., Amend A.,  
Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,  
Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,  
Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
Entian K.-D., Feibi H., Garcia-del Portillo F., Garrido P.,  
Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
Jones L.-M., Kaerst U., Krefz J., Kunz M., Kuntz F., Kurapkut G.,  
Machueno E., Maitouram A., Mata Vicente J., Ng E., Nedjari H.,  
Nordstiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,  
Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,  
Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,  
RT "Comparative genomics of Listeria species";  
RL Science 294:849-852(2001).  
DR EMBL; AL596166; CAC95955.1; -.  
DR PIR; AC1523; AC1523.  
DR Listlist; LIN0723; -.  
KW Complete proteome.  
SQ SEQUENCE 230 AA; 27050 MW; 05ACFPB3B3B871D CRC64;

Query Match 26.9%; Score 69.5; DB 2; Length 230;  
Best Local Similarity 30.5%; Pred. No. 12;

Matches 18; Conservative 10; Mismatches 16; Indels 15; Gaps 1;  
Qy 2 EKERPVKQ-----CQRLQIEMNELLNEVALQVDRKVADEEKQSYDAVV 45  
Db 33 EVESPYKELEQLNHOKELEKMSAIEIEQQLANKERQAIIEELKQSAEAI 91

RESULT 15  
ID Q9YDX9 PRELIMINARY; PRT; 297 AA.  
AC Q9YDX9;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein APE0790.  
GN OrderedLocustNames=APE0790;  
OS Aeropyrum pernix.  
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcates;  
OC Desulfurococcaceae; Aeropyrum.  
OX NCBI\_TaxID=56636;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=KI.  
RX MEDLINE=99310339; PubMed=10382966;  
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,  
Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anka A., Kosugi H.,  
Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,  
Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,  
Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,  
Nakamura Y., Nomura N., Sako Y., Kikuchi H.;  
RT "Complete genome sequence of an aerobic hyper-thermophilic  
crenarchaeon, Aeropyrum pernix KI.";  
RL DNA Res. 6:83-101(1999).  
DR EMBL; AP000060; BAA79768.1; -.  
DR PIR; H72670; H72670.  
DR HSSP; O15813; ID7M.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 297 AA; 35816 MW; CBBFF39610635B8 CRC64;

Query Match 26.4%; Score 68; DB 2; Length 297;  
Best Local Similarity 46.7%; Pred. No. 23;  
Matches 14; Conservative 7; Mismatches 9; Indels 0; Gaps 0;  
Job time : 95.246 secs  
Qy 12 RLQIEMNELLNEVALQVDRKVADEEKQSY 41  
Db 29 RLKERRKLLNEVALREKRSRKREKREY 58

Search completed: November 3, 2005, 22:03:06

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2005, 21:51:39 ; Search time 27.2086 Seconds  
(without alignments)  
145.410 Million cell updates/sec

Title: 09782816-52

Perfect score: 258  
Sequence: 1 GKEKTPVQKCORLQIENNEL.....ADEEKQSYDAVAVTISTAR 53

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*

2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*

3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*

4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*

5: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep:\*

6: /cgn2\_6/ptodata/1/1aa/6D.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	160	62.0	142	4	US-09-270-767-43370
2	70.5	27.3	1937	4	US-09-538-092-918
3	62.5	24.2	1939	4	US-09-538-092-915
4	62.5	24.2	1939	4	US-09-949-016-11104
5	61.5	23.8	1939	4	US-09-949-016-6925
6	61.5	23.8	1942	4	US-09-949-016-8135
7	59.5	23.1	1940	4	US-09-538-092-901
8	59.5	23.1	1963	4	US-09-949-016-8888
9	58	22.5	1248	2	US-09-080-897-2
10	58	22.5	1248	2	US-09-323-735-2
11	58	22.5	1255	3	US-09-080-897-4
12	58	22.5	1255	3	US-08-899-595-1
13	58	22.5	1255	3	US-09-323-735-4
14	58	22.5	1315	3	US-08-899-595-3
15	58	22.5	1413	3	US-09-252-991A-23627
16	57	22.1	203	4	US-09-711-164-436
17	57	22.1	233	5	PCT-US94-01149-14
18	57	22.1	304	5	PCT-US94-01149-55
19	57	22.1	304	5	PCT-US94-01149-57
20	57	22.1	585	2	US-08-453-848-11
21	57	22.1	585	3	US-09-169-027-11
22	57	22.1	586	2	US-08-453-848-19
23	57	22.1	586	2	US-09-169-027-19
24	57	22.1	589	2	US-08-453-848-13
25	57	22.1	589	2	US-09-169-027-13
26	57	22.1	592	2	US-08-453-848-17
27	57	22.1	592	3	US-09-169-027-17

28	57	22.1	1686	4	US-09-355-160D-2	Sequence 2, Appli
29	57	22.1	1686	4	US-10-092-219-2	Sequence 2, Appli
30	56.5	21.9	231	4	US-09-270-767-58576	Sequence 58576, A
31	56.5	21.9	442	4	US-09-270-767-43234	Sequence 43234, A
32	56	21.7	179	4	US-09-270-767-60196	Sequence 60196, A
33	56	21.7	592	4	US-09-919-039-141	Sequence 141, App
34	56	21.7	632	4	US-09-949-016-10668	Sequence 10668, A
35	56	21.7	661	4	US-09-914-259-66	Sequence 66, Appl
36	56	21.7	1572	4	US-09-562-702A-32	Sequence 32, Appl
37	56	21.7	1572	4	US-09-561-818A-28	Sequence 28, Appl
38	56	21.7	1605	4	US-09-562-702A-30	Sequence 30, Appl
39	56	21.7	1605	4	US-09-561-818A-26	Sequence 26, Appl
40	55.5	21.5	319	3	US-08-872-979-1	Sequence 1, Appli
41	55.5	21.5	953	4	US-09-252-991A-21559	Sequence 21559, A
42	55.5	21.5	1938	4	US-09-949-016-6417	Sequence 6417, Ap
43	55.5	21.5	1959	4	US-09-949-016-8134	Sequence 8134, Ap
44	55	21.3	182	4	US-09-248-796A-14143	Sequence 14143, A
45	55	21.3	208	4	US-09-248-796A-19287	Sequence 19287, A

#### ALIGNMENTS

```
RESULT 1
US-09-270-767-43370
; Sequence 43370, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 43370
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-43370

Query Match          62.0% Score 160; DB 4; Length 142;
Best Local Similarity 100.0%; Pred. No. 7.4e-15;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 GKEKTPVQKCORLQIENNELNEVAALQVDRK 32
Db 111 GKEKTPVQKCORLQIENNELNEVAALQVDRK 142

RESULT 2
US-09-538-092-918
; Sequence 918, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Gluc, Loic
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqFormatter Version 0.9
; SEQ ID NO 918
; LENGTH: 1937
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
```



```
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P13535
US-09-538-092-918

Query Match          27.3%; Score 70.5; DB 4; Length 1937;
Best Local Similarity 28.9%; Pred. No. 1.1;
Matches 22; Conservative 15; Mismatches 8; Indels 31; Gaps 4;

QY 2 EKRPVQKCO-----PLQIEMNLENEVALQVDRKVA--DEE-- 37
DB 1533 KQVVEKCEIQALMEKASLEHBEKILRIQLELNQKSE-----VDKRIAEKDESID 1587
QY 38 --KQSYDAVAVATVIST 51
DB 1588 QLKRNHTRVETMOST 1603

RESULT 3
US-09-538-092-915
; Sequence 915, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Gluc, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqFormatter Version 0.9
; SEQ ID NO 915
; LENGTH: 1939
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P12882
US-09-538-092-915

Query Match          24.2%; Score 62.5; DB 4; Length 1939;
Best Local Similarity 32.7%; Pred. No. 15;
Matches 16; Conservative 14; Mismatches 8; Indels 11; Gaps 2;

QY 9 KCORLQIEMNLENEVALQVDRKVA-----DEEKQSYDAVAVATVIST 51
DB 1561 KIIRIQLELNQKSE-----VDKRIAEKDESIDQKRNHTRVETMOST 1604

RESULT 4
US-09-949-016-11104
; Sequence 1104, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11104

; LENGTH: 1939
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11104

Query Match          24.2%; Score 62.5; DB 4; Length 1939;
Best Local Similarity 32.7%; Pred. No. 15;
Matches 16; Conservative 14; Mismatches 8; Indels 11; Gaps 2;

QY 9 KCORLQIEMNLENEVALQVDRKVA-----DEEKQSYDAVAVATVIST 51
DB 1561 KIIRIQLELNQKSE-----VDKRIAEKDESIDQKRNHTRVETMOST 1604

RESULT 5
US-09-949-016-6925
; Sequence 6925, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6925
; LENGTH: 1939
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6925

Query Match          23.8%; Score 61.5; DB 4; Length 1939;
Best Local Similarity 36.7%; Pred. No. 21;
Matches 18; Conservative 13; Mismatches 7; Indels 11; Gaps 3;

QY 9 KCORLQIEMNLENEVALQVDRKVA-----DEE--KQSYDAVAVATVIST 51
DB 1561 KIIRIQLELNQKSE-----IDKRIAEKDESIDQKRNHTRVETMOST 1604

RESULT 6
US-09-949-016-8135
; Sequence 8135, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8135
; LENGTH: 1942
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8135
```





ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 75 DENISE DRIVE  
CITY: HILLSBOROUGH  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94010  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/323,735  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/080,897  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: OSMAN, RICHARD A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: UW97-001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 343-4341  
TELEFAX: (650) 343-4342  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1248 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-323-735-2

Query Match 22.5%; Score 58; DB 3; Length 1248;  
Best Local Similarity 34.0%; Pred. No. 37;  
Matches 16; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

Oy 2 EKETPVOKCRLQIEMNELINVEVALQVDRKVADEKQSYDAVVA 48  
Db 517 EKQIATKQKDLBAVSKLTGEVAKLTSELEDAKNEMASISAVVAVSVSSA 563

RESULT 11  
US-09-080-897-4  
Sequence 4, Application US/09080897  
Patent No. 5985574  
GENERAL INFORMATION:  
APPLICANT: King, Mary-Claire  
APPLICANT: Lynch, Eric D.  
APPLICANT: Lee, Ming  
APPLICANT: Morrow, Jan E.  
APPLICANT: Welsh, Pili L.  
APPLICANT: Leon, Pedro E.  
TITLE OF INVENTION: Modulators of Actin  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 75 DENISE DRIVE  
CITY: HILLSBOROUGH  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94010  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/080,897  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: OSMAN, RICHARD A

REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: UW97-001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 343-4341  
TELEFAX: (650) 343-4342  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-080-897-4

Query Match 22.5%; Score 58; DB 2; Length 1255;  
Best Local Similarity 35.8%; Pred. No. 37;  
Matches 19; Conservative 10; Mismatches 22; Indels 2; Gaps 1;

Oy 2 EKETPVOKCRLQIEMNELINVEVALQVDRKVADEKQSYDAVVA--TWISTA 52  
Db 517 EKQIATKQKDLBAVSKLTGEVAKLTSELEDAKNEMASISAVVAVSVSSA 569

RESULT 12  
US-08-899-595-1  
Sequence 1, Application US/08895595  
Patent No. 611072  
GENERAL INFORMATION:  
APPLICANT: Natumiya, Shuh  
APPLICANT: Takahashi, No. 611072uaki  
TITLE OF INVENTION: RHO TARGET PROTEIN HUMAN MDIA AND GENE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/899,595  
FILING DATE: 24-JUL-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 8-242701  
FILING DATE: 26-AUG-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 9-90170  
FILING DATE: 25-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Stephen A. Bent  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 049441/0112  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 672-5300  
TELEFAX: (202) 672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-899-595-1

Matches 19; Conservative 10; Mismatches 22; Indels 2; Gaps 1;

QY 2 EKETPVQKQRLQIEMNELNEVALQVDRVADBEKQSYAVVA--TVISTA 52  
Db 517 EKQOITAKQDLEAVSKLTGEVAKLTKELEDAKKEMASLSAVVAVPSVSSA 569

## RESULT 13

US-09-323-735-4  
Sequence 4, Application US/09323735  
Patent No. 6197932  
GENERAL INFORMATION:  
APPLICANT: King, Mary-Claire  
APPLICANT: Lynch, Eric D.  
APPLICANT: Lee, Ming  
APPLICANT: Morrow, Jan E.  
APPLICANT: Welch, Piri L.  
APPLICANT: Leon, Pedro E.  
TITLE OF INVENTION: Modulators of Actin  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 75 DENISE DRIVE  
CITY: HILLSBOROUGH  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94010  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/323,735  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/080,897  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: OSMAN, RICHARD A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: UM97-001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 343-4341  
TELEFAX: (650) 343-4342  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-323-735-4

Query Match 22.5%; Score 58; DB 3; Length 1255;  
Best Local Similarity 35.8%; Pred. No. 37;  
Matches 19; Conservative 10; Mismatches 22; Indels 2; Gaps 1;

QY 2 EKETPVQKQRLQIEMNELNEVALQVDRVADBEKQSYAVVA--TVISTA 52  
Db 517 EKQOITAKQDLEAVSKLTGEVAKLTKELEDAKKEMASLSAVVAVPSVSSA 569

## RESULT 14

US-08-899-595-3  
Sequence 3, Application US/08899595  
Patent No. 611072  
GENERAL INFORMATION:  
APPLICANT: Narumiya, Shuh  
APPLICANT: Takahashi, No. 611072uaki  
TITLE OF INVENTION: RHO TARGET PROTEIN HUMAN MDIA AND GENE  
TITLE OF INVENTION: ENCODING SAME

## NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/899,595  
FILING DATE: 24-JUL-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 8-242701  
FILING DATE: 26-AUG-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 9-90170  
FILING DATE: 25-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Stephen A. Bent  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 049441/0112  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 672-5300  
TELEFAX: (202) 672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1315 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-899-595-3

Query Match 22.5%; Score 58; DB 3; Length 1315;  
Best Local Similarity 34.0%; Pred. No. 39;  
Matches 16; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY 2 EKETPVQKQRLQIEMNELNEVALQVDRVADBEKQSYAVVATV 48  
Db 569 EKQOITAKQDLEAVSQLTGEVAKLTKELEDAKKEMASLSAAITV 615

## RESULT 15

US-09-252-991A-23627  
Sequence 23627, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
PRIOR FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 23627  
LENGTH: 1413  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-23627

Query Match 22.5%; Score 58; DB 4; Length 1413;  
Best Local Similarity 31.5%; Pred. No. 43;  
Matches 17; Conservative 11; Mismatches 16; Indels 10; Gaps 2;



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2005, 21:53:15 ; Search time 95.5134 Seconds  
(without alignments)  
232.174 Million cell updates/sec

Title: 09782816-52

Perfect score: 258  
Sequence: 1 GKEKTPYQKCRLOIEMNEL.....ADERKQSYDAVAVATISTAR 53

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1667879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1667879

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US10F\_PUBCOMB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	258	100.0	53	10	US-09-782-816A-52
2	243.5	94.4	380	10	US-09-782-816A-56
3	243.5	94.4	380	20	US-11-097-143-4056
4	114	44.2	22	10	US-09-782-816A-26
5	108	41.9	21	10	US-09-782-816A-27
6	105	40.7	22	10	US-09-782-816A-2
7	103	39.9	20	10	US-09-782-816A-28
8	98	38.0	19	10	US-09-782-816A-29
9	93	36.0	18	10	US-09-782-816A-30
10	88	34.1	17	10	US-09-782-816A-31
11	81	31.4	16	10	US-09-782-816A-32

	77	29.8	15	10	US-09-782-816A-33	Sequence 33, Appl
12	74	28.7	183	10	US-09-782-816A-54	Sequence 54, Appl
13	74	28.7	224	16	US-10-425-115-315631	Sequence 315631,
14	72.5	28.1	52	10	US-09-782-816A-51	Sequence 51, Appl
15	72.5	28.1	401	16	US-10-408-765A-1369	Sequence 1369, Ap
16	72.5	28.1	406	10	US-09-782-816A-53	Sequence 53, Appl
17	72.5	28.1	465	9	US-09-925-298-676	Sequence 676, App
18	72.5	28.1	465	14	US-10-102-806-676	Sequence 676, App
19	72.5	28.1	14	10	US-09-782-816A-34	Sequence 34, Appl
20	72	27.9	14	10	US-10-723-860-106	Sequence 106, App
21	70.5	27.3	1937	16	US-10-723-860-106	Sequence 106, App
22	70.5	27.3	1937	16	US-10-723-860-106	Sequence 106, App
23	69.5	26.9	97	9	US-09-864-761-43401	Sequence 43401, A
24	67	26.0	13	17	US-09-782-816A-35	Sequence 35, Appl
25	64	24.8	886	17	US-10-732-923-3311	Sequence 3311, Ap
26	63.5	24.6	185	15	US-10-424-599-207072	Sequence 207072,
27	63.5	24.6	233	15	US-10-425-114-48404	Sequence 48404, A
28	63.5	24.6	324	15	US-10-424-599-207069	Sequence 207069,
29	62.5	24.2	97	9	US-09-864-761-39043	Sequence 39043, A
30	62.5	24.2	103	9	US-09-864-761-42193	Sequence 42193, A
31	62.5	24.2	141	17	US-10-805-684-134	Sequence 134, App
32	62.5	24.2	1939	16	US-10-205-219-102	Sequence 102, App
33	62.5	24.2	1939	16	US-10-408-765A-2188	Sequence 2188, App
34	62.5	24.2	1939	17	US-10-805-684-152	Sequence 152, App
35	62	24.0	351	16	US-10-821-273-26	Sequence 26, Appl
36	61.5	23.8	97	9	US-09-864-761-42176	Sequence 42176, A
37	61.5	23.8	483	15	US-10-424-599-170198	Sequence 170198,
38	61.5	23.8	1948	18	US-10-450-763-51592	Sequence 51592, A
39	61	23.6	1985	17	US-10-732-923-3351	Sequence 3351, Ap
40	60	23.3	245	16	US-10-767-701-41002	Sequence 41002, A
41	60	23.3	1042	16	US-10-408-765A-1655	Sequence 1655, App
42	59.5	23.1	1167	16	US-10-723-860-54	Sequence 54, Appl
43	59.5	23.1	1940	10	US-09-738-630-99	Sequence 99, Appl
44	59.5	23.1	1940	16	US-10-408-765A-1175	Sequence 1175, Appl
45	59	22.9	110	15	US-10-335-977-8524	Sequence 8524, Ap

#### ALIGNMENTS

RESULT 1  
US-09-782-816A-52  
Sequence 52, Application US/09782816A  
Publication No. US20030032771A1  
GENERAL INFORMATION:  
APPLICANT: Sharp, David J.  
APPLICANT: Rogers, Gregory C.  
APPLICANT: Scholey, Jonathan M.  
TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR  
FILE REFERENCE: UC069, 001A  
CURRENT APPLICATION NUMBER: US/09/782, 816A  
CURRENT FILING DATE: 2001-02-14  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 52  
LENGTH: 53  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-09-782-816A-52

Query Match 100.0%; Score 258; DB 10; Length 53;  
Best Local Similarity 100.0%; Pred. No. 1.9e-25;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKEKTPYQKCRLOIEMNELNEVAALQVDRKVDDEKQSYDAVAVATISTAR 53  
DB 1 GKEKTPYQKCRLOIEMNELNEVAALQVDRKVDDEKQSYDAVAVATISTAR 53

RESULT 2  
US-09-782-816A-56  
Sequence 56, Application US/09782816A  
Publication No. US20030032771A1

```
/ GENERAL INFORMATION:
/ APPLICANT: Sharp, David J.
/ APPLICANT: Rogers, Gregory C.
/ APPLICANT: Scholey, Jonathan M.
/ TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
/ FILE REFERENCE: UC069.001A
/ CURRENT FILING DATE: 2001-02-14
/ NUMBER OF SEQ ID NOS: 56
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 56
/ LENGTH: 380
/ TYPE: PRT
/ ORGANISM: Drosophila melanogaster
US-09-782-816A-56
```

```
Query Match          94.4%; Score 243.5; DB 10; Length 380;
Best Local Similarity 98.1%; Pred. No. 1.6e-22;
Matches 52; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

```
Oy      1 GKEKTPVQKQRLQIEMNELINEVALQVDRKVADEEKOSYDAVAVATISTAR 53
Db      94 GKEKTPVQKQRLQIEMNELINEVALQVDRKVADEEKOSYDA-VATVISTAR 145
```

```
RESULT 3
US-11-097-143-4056
```

```
/ Sequence 4056, Application US/11097143
/ Publication No. US20050208558A1
/ GENERAL INFORMATION:
/ APPLICANT: Venter, J. Craig
/ APPLICANT: et al.
/ TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
/ TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
/ FILE REFERENCE: CL000728
/ CURRENT FILING DATE: 2005-04-04
/ PRIOR APPLICATION NUMBER: US/11/097,143
/ PRIOR FILING DATE: 1999-10-05
/ PRIOR APPLICATION NUMBER: 60/157,832
/ PRIOR FILING DATE: 1999-10-19
/ PRIOR APPLICATION NUMBER: 60/160,191
/ PRIOR FILING DATE: 1999-10-28
/ PRIOR APPLICATION NUMBER: 60/164,769
/ PRIOR FILING DATE: 1999-11-12
/ PRIOR APPLICATION NUMBER: 60/173,383
/ PRIOR FILING DATE: 1999-12-28
/ PRIOR APPLICATION NUMBER: 60/175,693
/ PRIOR FILING DATE: 2000-01-12
/ PRIOR APPLICATION NUMBER: 60/184,831
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: 60/191,637
/ PRIOR FILING DATE: 2000-03-23
/ NUMBER OF SEQ ID NOS: 43008
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4056
/ LENGTH: 380
/ TYPE: PRT
/ ORGANISM: DROSOPHILA
US-11-097-143-4056
```

```
Query Match          94.4%; Score 243.5; DB 20; Length 380;
Best Local Similarity 98.1%; Pred. No. 1.6e-22;
Matches 52; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

```
Oy      1 GKEKTPVQKQRLQIEMNELINEVALQVDRKVADEEKOSYDAVAVATISTAR 53
Db      94 GKEKTPVQKQRLQIEMNELINEVALQVDRKVADEEKOSYDA-VATVISTAR 145
```

```
RESULT 4
```

```
US-09-782-816A-26
/ Sequence 26, Application US/09782816A
/ Publication No. US20030032771A1
/ GENERAL INFORMATION:
/ APPLICANT: Sharp, David J.
/ APPLICANT: Rogers, Gregory C.
/ APPLICANT: Scholey, Jonathan M.
/ TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
/ FILE REFERENCE: UC069.001A
/ CURRENT FILING DATE: 2001-02-14
/ NUMBER OF SEQ ID NOS: 56
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 26
/ LENGTH: 22
/ TYPE: PRT
/ ORGANISM: Drosophila melanogaster
US-09-782-816A-26
```

```
Query Match          44.2%; Score 114; DB 10; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      1 GKEKTPVQKQRLQIEMNELIN 22
Db      1 GKEKTPVQKQRLQIEMNELIN 22
```

```
RESULT 5
US-09-782-816A-27
```

```
/ Sequence 27, Application US/09782816A
/ Publication No. US20030032771A1
/ GENERAL INFORMATION:
/ APPLICANT: Sharp, David J.
/ APPLICANT: Rogers, Gregory C.
/ APPLICANT: Scholey, Jonathan M.
/ TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
/ TITLE OF INVENTION: PROLIFERATION
/ FILE REFERENCE: UC069.001A
/ CURRENT FILING DATE: US/09/782,816A
/ CURRENT FILING DATE: 2001-02-14
/ NUMBER OF SEQ ID NOS: 56
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 27
/ LENGTH: 21
/ TYPE: PRT
/ ORGANISM: Drosophila melanogaster
US-09-782-816A-27
```

```
Query Match          41.9%; Score 108; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 9.5e-07;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      2 EKETPVQKQRLQIEMNELIN 22
Db      1 EKETPVQKQRLQIEMNELIN 21
```

```
RESULT 6
US-09-782-816A-2
```

```
/ Sequence 2, Application US/09782816A
/ Publication No. US20030032771A1
/ GENERAL INFORMATION:
/ APPLICANT: Sharp, David J.
/ APPLICANT: Rogers, Gregory C.
/ APPLICANT: Scholey, Jonathan M.
/ TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
/ TITLE OF INVENTION: PROLIFERATION
/ FILE REFERENCE: UC069.001A
/ CURRENT FILING DATE: US/09/782,816A
/ CURRENT FILING DATE: 2001-02-14
/ NUMBER OF SEQ ID NOS: 56
```

```
SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-782-816A-2

Query Match      40.7%; Score 105; DB 10; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      23 EVAALQVDRKVADEKOSYDAV 44
Db      1 EVAALQVDRKVADEKOSYDAV 22

RESULT 7
US-09-782-816A-28
; Sequence 28, Application US/09782816A
; Publication No. US20030032771A1
; GENERAL INFORMATION:
; APPLICANT: Sharp, David J.
; APPLICANT: Rogers, Gregory C.
; APPLICANT: Scholey, Jonathan M.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
; FILE REFERENCE: UC069.001A
; CURRENT APPLICATION NUMBER: US/09/782,816A
; CURRENT FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-782-816A-28

Query Match      39.9%; Score 103; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.9e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 KETPVQKCRLOIEMNELLN 22
Db      1 KETPVQKCRLOIEMNELLN 20

RESULT 8
US-09-782-816A-29
; Sequence 29, Application US/09782816A
; Publication No. US20030032771A1
; GENERAL INFORMATION:
; APPLICANT: Sharp, David J.
; APPLICANT: Rogers, Gregory C.
; APPLICANT: Scholey, Jonathan M.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
; FILE REFERENCE: UC069.001A
; CURRENT APPLICATION NUMBER: US/09/782,816A
; CURRENT FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-782-816A-29

Query Match      38.0%; Score 98; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 ETPVQKCRLOIEMNELLN 22
Db      1 ETPVQKCRLOIEMNELLN 22
```

```
Db      1 ETPVQKCRLOIEMNELLN 19

RESULT 9
US-09-782-816A-30
; Sequence 30, Application US/09782816A
; Publication No. US20030032771A1
; GENERAL INFORMATION:
; APPLICANT: Sharp, David J.
; APPLICANT: Rogers, Gregory C.
; APPLICANT: Scholey, Jonathan M.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
; FILE REFERENCE: UC069.001A
; CURRENT APPLICATION NUMBER: US/09/782,816A
; CURRENT FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-782-816A-30

Query Match      36.0%; Score 93; DB 10; Length 18;
Best Local Similarity 100.0%; Pred. No. 6.5e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 TPVQKCRLOIEMNELLN 22
Db      1 TPVQKCRLOIEMNELLN 18

RESULT 10
US-09-782-816A-31
; Sequence 31, Application US/09782816A
; Publication No. US20030032771A1
; GENERAL INFORMATION:
; APPLICANT: Sharp, David J.
; APPLICANT: Rogers, Gregory C.
; APPLICANT: Scholey, Jonathan M.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
; FILE REFERENCE: UC069.001A
; CURRENT APPLICATION NUMBER: US/09/782,816A
; CURRENT FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-782-816A-31

Query Match      34.1%; Score 88; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 PVQKCRLOIEMNELLN 22
Db      1 PVQKCRLOIEMNELLN 17

RESULT 11
US-09-782-816A-32
; Sequence 32, Application US/09782816A
; Publication No. US20030032771A1
; GENERAL INFORMATION:
; APPLICANT: Sharp, David J.
; APPLICANT: Rogers, Gregory C.
; APPLICANT: Scholey, Jonathan M.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
```

FILE REFERENCE: UC069.001A  
CURRENT APPLICATION NUMBER: US/09/782.816A  
CURRENT FILING DATE: 2001-02-14  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 32  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-09-782-816A-32

Query Match 31.4%; Score 81; DB 10; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.0019;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 VQKQRLQIEMNELN 22  
DB 1 VQKQRLQIEMNELN 16

RESULT 12  
US-09-782-816A-33  
Sequence 33, Application US/09782816A  
Publication No. US20030032771A1  
GENERAL INFORMATION:  
APPLICANT: Sharp, David J.  
APPLICANT: Rogers, Gregory C.  
APPLICANT: Scholey, Jonathan M.  
TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR  
TITLE OF INVENTION: PROLIFERATION  
FILE REFERENCE: UC069.001A  
CURRENT APPLICATION NUMBER: US/09/782.816A  
CURRENT FILING DATE: 2001-02-14  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 33  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-09-782-816A-33

Query Match 29.8%; Score 77; DB 10; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0058;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 OKCQRLQIEMNELN 22  
DB 1 OKCQRLQIEMNELN 15

RESULT 13  
US-09-782-816A-54  
Sequence 54, Application US/09782816A  
Publication No. US20030032771A1  
GENERAL INFORMATION:  
APPLICANT: Sharp, David J.  
APPLICANT: Rogers, Gregory C.  
APPLICANT: Scholey, Jonathan M.  
TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR  
TITLE OF INVENTION: PROLIFERATION  
FILE REFERENCE: UC069.001A  
CURRENT APPLICATION NUMBER: US/09/782.816A  
CURRENT FILING DATE: 2001-02-14  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 54  
LENGTH: 183  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-782-816A-54

Query Match 28.7%; Score 74; DB 10; Length 183;  
Best Local Similarity 35.8%; Pred. No. 0.32;

Matches 19; Conservative 10; Mismatches 24; Indels 0; Gaps 0;  
QY 1 GKEKTPVQKQRLQIEMNELNEVAALQVDRKVADEKOSYDAVATVISTAR 53  
DB 94 GVKETPOOKYORLHVEVOELTTEVEKIKITVKSATEKULPVVALQALAK 146

RESULT 14  
US-10-425-115-315831  
Sequence 315831, Application US/10425115  
Publication No. US20040214272A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/425.115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 315831  
LENGTH: 224  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_51109C.1.dep  
US-10-425-115-315831

Query Match 28.7%; Score 74; DB 16; Length 224;  
Best Local Similarity 35.8%; Pred. No. 0.41;  
Matches 19; Conservative 10; Mismatches 24; Indels 0; Gaps 0;

QY 1 GKEKTPVQKQRLQIEMNELNEVAALQVDRKVADEKOSYDAVATVISTAR 53  
DB 47 GVKETPOOKYORLHVEVOELTTEVEKIKITVKSATEKULPVVALQALAK 99

RESULT 15  
US-09-782-816A-51  
Sequence 51, Application US/09782816A  
Publication No. US20030032771A1  
GENERAL INFORMATION:  
APPLICANT: Sharp, David J.  
APPLICANT: Rogers, Gregory C.  
APPLICANT: Scholey, Jonathan M.  
TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR  
TITLE OF INVENTION: PROLIFERATION  
FILE REFERENCE: UC069.001A  
CURRENT APPLICATION NUMBER: US/09/782.816A  
CURRENT FILING DATE: 2001-02-14  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 51  
LENGTH: 52  
TYPE: PRT  
ORGANISM: Unknown  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: 44  
OTHER INFORMATION: Xaa = Val or Leu  
OTHER INFORMATION: A sequence conserved among Homo sapiens and Mus  
OTHER INFORMATION: musculus.  
US-09-782-816A-51

Query Match 28.1%; Score 72.5; DB 10; Length 52;  
Best Local Similarity 48.7%; Pred. No. 0.1;  
Matches 19; Conservative 5; Mismatches 14; Indels 1; Gaps 1;

QY 1 GKEKTPVQKQRLQIEMNELNEVAALQVDRKVADEK 38  
DB 1 GVKETPOOKYORLHVEVOELTTEVEKIKITVKSATEK 39



Search completed: November 3, 2005, 22:11:41  
Job time : 96.5134 secs

09782816-52.rapb

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: November 3, 2005, 21:43:09 ; Search time 18.9091 Seconds  
(without alignments)  
264.596 Million cell updates/sec

Title: 09782816-51

Sequence: 1 GKETPOQKQRLHVEVQL.....ESATBEKLTPLVLAQAL 52

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69.5	27.7	804	2	TJ2864
2	68	27.1	1742	2	T49451
3	67.5	26.9	1325	2	S16129
4	67	26.7	1281	2	TJ2235
5	67	26.7	2481	2	D80011
6	66	26.3	393	2	T04957
7	66	26.3	866	2	C71509
8	65.5	26.1	1837	2	T41023
9	64	25.5	210	2	B84499
10	63.5	25.3	163	2	S69518
11	63.5	25.3	374	2	D70391
12	63	25.1	281	2	E75216
13	63	25.1	314	2	P90577
14	62.5	24.9	1281	2	JCS368
15	62	24.7	296	2	TJ2222
16	62	24.7	2469	2	H36812
17	61	24.3	442	2	C95070
18	61	24.3	442	2	A97938
19	60.5	24.1	819	2	AC2029
20	60.5	24.1	1214	2	JC2069
21	60	23.9	523	2	S06920
22	60	23.9	695	2	H72243
23	59.5	23.7	279	2	D71453
24	59.5	23.7	665	2	S62328
25	59	23.5	112	2	G72645
26	59	23.5	239	2	D84004
27	59	23.5	550	2	G85436
28	59	23.5	611	2	E90543
29	59	23.5	782	2	A34219

30	59	23.5	951	2	T45726	hypothetical prote
31	59	23.5	2101	2	A42184	nuclear mitotic ap
32	58.5	23.3	154	2	T25602	hypothetical prote
33	58.5	23.3	821	2	A12417	hypothetical prote
34	58	23.1	302	2	T01615	hypothetical prote
35	58	23.1	319	2	T25666	hypothetical prote
36	58	23.1	623	2	A48315	hypothetical prote
37	58	23.1	630	2	D97992	lamin IIR - Africa
38	58	23.1	677	2	T22333	hypothetical prote
39	58	23.1	707	2	S78538	histidine ammonia-
40	58	23.1	727	2	AD2188	site-specific reco
41	58	23.1	740	2	G95153	hypothetical prote
42	58	23.1	978	2	A70387	neuraminidase, pro
43	58	23.1	2139	2	T18286	conserved hypother
44	58	23.1	2422	2	T12687	myosin heavy chain
45	57.5	22.9	237	2	C87656	AKR protein homolo
						GGDEF family prote

## ALIGNMENTS

```

RESULT 1
TJ2864
hypothetical protein ZK484.4 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C/Accession: TJ2864
R/Murray, J.; Mohlmann, P.
submitted to the EMBL Data Library, December 1997
A/Description: The sequence of C. elegans cosmid ZK484.
A/Reference number: Z21237
A/Accession: TJ2864
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-804 <MB>
A/Cross-references: UNIPROT:Q95X56; EMBL:AF040659; PDB:AA095062.1; GSPDB:GN00019; CESP:
A/Experimental source: strain Bristol N2; clone ZK484
C/Genetics:
A/Gene: CESP:ZK484.4
A/Map position: 1
A/Introns: 52/1; 113/1; 162/3; 201/3; 248/2; 346/1; 433/2; 464/3; 565/3; 628/2; 646/3; 7
C/Superfamily: Caenorhabditis elegans hypothetical protein ZK484.4

Query Match 27.7%; Score 69.5; DB 2; Length 804;
Best local similarity 34.4%; Pred. No. 20;
Matches 22; Conservative 10; Mismatches 19; Indels 13; Gaps 3;

Qy 2 VKETPOQKQRLHVEVQLTTEV-----KTKT-TVKSATBE-KLTPVLAQ 48
Db 366 VEDTPDEKYQKTKTVSEIENNEEDGIGASTSKRTVTIKEDIEVROTPAKLRA 425

Oy 49 LAAL 52
Db 426 SAKL 429

RESULT 2
T49451
kinesin-like protein Kif2a related protein [imported] - Neurospora crassa
N/Alternate names: protein B14D6.30
C/Species: Neurospora crassa
C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
R/Schulte, U.; Aign, V.; Hohnsbeil, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A/Reference number: Z25022
A/Accession: T49451
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1742 <SCH>
A/Cross-references: UNIPROT:Q8X0C5; EMBL:AL356173; GSPDB:GN00116; NCSP:B14D6.30
A/Experimental source: BAC clone B14D6; strain OR74A
C/Genetics:

```

A:Gene: NCSP:BI4D6.30  
A:Map position: 6  
A:Introns: 58/1; 166/1; 267/3; 1543/3

Query Match 27.1%; Score 68; DB 2; Length 1742;  
Best Local Similarity 36.6%; Pred. No. 66;  
Matches 15; Conservative 8; Mismatches 18; Indels 0; Gaps 0;

Oy 7 OOKYORLHAEVOELTTEVEKIKTKTKESATEBKLTPLVLAKL 47  
Db 950 EKKHQTLDMEVELKTEIAKQALSVESISRTSTPVLIRK 990

RESULT 3  
S16129  
dysenin-associated protein, 150K, cytosolic - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S16129  
R:Holzbaur, E.L.F.; Hammarback, J.A.; Paschal, B.M.; Kravitt, N.G.; Pfister, K.K.; Vallee  
Nature 351, 579-583, 1991  
A:Title: Homology of a 150K cytoplasmic dynein-associated polypeptide with the Drosophila  
A:Reference number: S16129; MUID:91260877; PMID:1828535  
A:Accession: S16129  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1325 <KKV>  
A:Cross-references: UNIPROT:P28023

Query Match 26.9%; Score 67.5; DB 2; Length 1325;  
Best Local Similarity 34.7%; Pred. No. 56;  
Matches 17; Conservative 11; Mismatches 20; Indels 1; Gaps 1;

Oy 1 GVKETPOOKYORLHAEVOELTTEVEKIKTKTKESATEBKLTPLVLAKL 49  
Db 314 GAESLQOEVALKRVDELTTDLTKAEIEKSGDGAASYQL-KQL 361

RESULT 4  
T22235  
hypothetical protein F45G2.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T22235  
R:Lindsay, S.  
submitted to the EMBL Data Library, March 1997  
A:Reference number: Z19535  
A:Accession: T22235  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1287 <WLL>  
A:Cross-references: UNIPROT:O62245; EMBL:Z93382; PIDD:CA807612.1; GSPDB:GN00021; CESP:F4  
A:Experimental source: clone F45G2  
C:Genetics:  
A:Gene: CESP:F45G2.3  
A:Map position: 3  
A:Introns: 59/2; 117/3; 153/3; 180/2; 233/2; 341/3; 393/1; 424/2; 551/3; 664/3; 734/3; 8

Query Match 26.7%; Score 67; DB 2; Length 1287;  
Best Local Similarity 31.6%; Pred. No. 61;  
Matches 18; Conservative 11; Mismatches 18; Indels 10; Gaps 2;

Oy 4 ETPOOKYORLHAEVOELTTEVEKIKTKTKESATEBKLTPLVLAKL 50  
Db 1122 EPPPEKOKILVEVDIDDFSEFQEKVKENTKESVAEEKVENSNEELKPPILKRAVS 1178

RESULT 5  
D90011  
FmbB protein (imported) - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: D90011

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguci  
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; I  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.  
A:Reference number: A89758; MUID:21311952; PMID:11418146  
A:Accession: D90011  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-2481 <KUR>  
A:Cross-references: UNIPROT:Q99QRE; GB:BA000018; PID:G13701961; PIDD:BA843253.1; GSPDB:GT  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: fmbB(mrp)

Query Match 26.7%; Score 67; DB 2; Length 2481;  
Best Local Similarity 35.1%; Pred. No. 1.2e+02;  
Matches 13; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

Oy 2 VKETPOOKYORLHAEVOELTTEVEKIKTKTKESATEE 38  
Db 1609 IEQTPNMSQGEINDAKQEVDTLNQAKTNDQSTNE 1645

RESULT 6  
T04957  
hypothetical protein F7J7.200 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: T04957  
R:Beyan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X  
submitted to the Protein Sequence Database, July 1998  
A:Reference number: Z15391  
A:Accession: T04957  
A:Molecule type: DNA  
A:Residues: 1-393 <BEV>  
A:Cross-references: UNIPROT:O49567; EMBL:AL021960  
A:Experimental source: cultivar Columbia; BAC clone F7J7  
C:Genetics:  
A:Map position: 4  
A:Introns: 80/1; 217/3; 264/1; 291/3  
A:Note: F7J7.200

Query Match 26.3%; Score 66; DB 2; Length 393;  
Best Local Similarity 38.6%; Pred. No. 21;  
Matches 17; Conservative 12; Mismatches 7; Indels 8; Gaps 2;

Oy 3 KETPOOKYORLHAEVOELTTEVEKIKTKTKESATEE-KLTPLVL 45  
Db 27 QESTQOTHQNLVYKVQ-----KWRSLKDSDAELKLSPALV 63

RESULT 7  
C71509  
probable DNA polymerase I - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
C:Species: Chlamydia trachomatis  
C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 09-Jul-2004  
C:Accession: C71509  
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marache, R.; Aravind, L.; Mitchell,  
Science 282, 754-759, 1998  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia traci*  
A:Reference number: A71570; MUID:9900809; PMID:9784136  
A:Accession: C71509  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-866 <ARN>  
A:Cross-references: UNIPROT:O84500; GB:AE001322; GB:AE001273; NID:G3328916; PIDD:AAC6809  
A:Experimental source: serotype D, strain UW-3/Cx  
C:Genetics:  
A:Gene: polA  
C:Superfamily: DNA-directed DNA polymerase I

Query Match 26.3%; Score 66; DB 2; Length 866;

Best Local Similarity 36.4%; Pred. No. 50;  
Matches 20; Conservative 10; Mismatches 21; Indels 4; Gaps 2;  
QY 1 GVKETPQOKYRLHVE---VOELTEVEKIKTTVESATEEKLTPVLLAKQIAL 52  
Db 188 GVSQGGPKKAQALTKRFQSVSELVANTRLSGKTQMIEDQKET-TLLSKRLATL 241

RESULT 8  
T41023  
probable nuclear pore complex-associated protein - fission yeast (Schizosaccharomyces po  
C/Species: Schizosaccharomyces pombe  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C/Accession: T41023  
R/Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, June 1998  
A/Accession: T41023  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-1837 <MUR>  
A/Cross-references: UNIPROT:O74424; EMBL:AL023860; PIDN:CAA19588.1; GSPDB:GN00068; SPDB:  
A/Experimental source: strain 972h-; cosmid c162  
C/Genetics:  
A/Gene: SPDB:SPC162.08C  
A/Map position: 3

Query Match 26.1%; Score 65.5; DB 2; Length 1837;  
Best Local Similarity 35.8%; Pred. No. 1.3e+02;  
Matches 19; Conservative 10; Mismatches 17; Indels 7; Gaps 2;  
QY 7 QOKYQRLHVEQELTEVEKIKTTV--KESATEEKLTPVLLAK-----QLAAL 52  
Db 662 QTSYQSRSLRIQLQTLNLESLKSISSRNKKFEALISSLOLEKSNIQQLTSL 714

RESULT 9  
E84499  
hypothetical protein At2g11890 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 12-Jul-2004  
C/Accession: E84499  
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Unayam, L.; Tallon, L.  
euser, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: AB4420; MUID:20083487; PMID:10617197  
A/Accession: E84499  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-210 <STO>  
A/Cross-references: UNIPROT:Q9S1Y3; GB:AB02093; NID:g4557062; PIDN:AA022501.1; GSPDB:GN  
C/Genetics:  
A/Gene: At2g11890  
A/Map position: 2  
C/Superfamily: uncharacterized CYTH domain protein

Query Match 25.5%; Score 64; DB 2; Length 210;  
Best Local Similarity 40.5%; Pred. No. 17;  
Matches 17; Conservative 9; Mismatches 14; Indels 2; Gaps 2;  
QY 1 GVK-ETPQOKYQ-RLLHVEQELTEVEKIKTTVESATEEKL 40  
Db 150 GVKLEVDETKYDFGNCYRIECETEPERVKTMIEEFLTEKTI 191

RESULT 10  
S69518  
hypothetical protein 12 - phage HPI  
C/Species: phage HPI  
C/Date: 06-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C/Accession: S69518

R. Esposito, D.; Fitzmaurice, W.P.; Benjamin, R.C.; Goodman, S.D.; Waldman, A.S.; Socca,  
Nucleic Acids Res. 24, 2360-2368, 1996  
A/Title: The complete nucleotide sequence of bacteriophage HPI DNA.  
A/Reference number: S69503; MUID:96279738; PMID:8710508  
A/Accession: S69518  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-163 <ESP>  
A/Cross-references: UNIPROT:P51714; EMBL:U24159; NID:g1046235; PIDN:AA09197.1; PID:g104  
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995  
C/Superfamily: phage HPI hypothetical protein 12

Query Match 25.3%; Score 63.5; DB 2; Length 163;  
Best Local Similarity 37.8%; Pred. No. 15;  
Matches 17; Conservative 11; Mismatches 16; Indels 1; Gaps 1;  
QY 5 TPQOKYQRLHVEQELTEVEKIK-TTVKESATEEKLTPVLLAKQ 48  
Db 25 TPQOKYQRLHVEQELTEVEKIK-TTVKESATEEKLTPVLLAKQ 69

RESULT 11  
D70391  
hypothetical protein aq\_1060 - Aquifex aeolicus  
C/Species: Aquifex aeolicus  
C/Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
C/Accession: D70391  
R/Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov  
V.  
Nature 392, 353-358, 1998  
A/Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A/Reference number: A70300; MUID:98196666; PMID:9537320  
A/Accession: D70391  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-374 <AOP>  
A/Cross-references: UNIPROT:O67159; GB:AB000721; NID:g2983544; PIDN:AAC07129.1; PID:g298  
A/Experimental source: strain VFS  
C/Genetics:  
A/Gene: aq\_1060  
C/Superfamily: multidrug resistance protein A; lipoyl/biotin-binding homology

Query Match 25.3%; Score 63.5; DB 2; Length 374;  
Best Local Similarity 31.1%; Pred. No. 36;  
Matches 19; Conservative 14; Mismatches 13; Indels 15; Gaps 3;  
QY 6 PQOKYQ-----LHVE-----VOELTEVEKIKTTVESATEEKLTPVLLAKQIAL 51  
Db 164 PRRFEEVDTNLKVLLHREYLEKSIQEIINTEIRAKKGI-ENARNEKTEIEELKEKLS 222

QY 52 L 52  
Db 223 L 223

RESULT 12  
F75216  
hypothetical protein PAB2181 - Pyrococcus abyssi (strain Orsay)  
C/Species: Pyrococcus abyssi  
C/Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C/Accession: F75216  
R/anonymous, Genoscope  
submitted to the EMBL Data Library, July 1999  
A/Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru  
A/Reference number: A75001  
A/Accession: F75216  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-281 <KAM>  
A/Cross-references: UNIPROT:Q9V217; GB:A0248283; GB:AL096836; NID:g5457433; PIDN:CAB4918  
A/Experimental source: strain Orsay  
C/Genetics:  
A/Gene: PAB2181

Query Match	25.1%;	Score 63;	DB 2;	Length 281;
Best Local Similarity	36.6%;	Pred. No. 30;		
Matches 15;	Conservative 12;	Mismatches 12;	Indels 2;	Gaps 2

```

QY      1 GYKE -TPQOKQRLIHVQELTTEVEKIKITVKESATHEKL 40
      ||:::||::||::||::||
DB      119 GIKEVAREFEYKLLKEYEKLQEFBEVKAKI -EAELESL 156

```

```

RESULT 13
P90577
1.lipidprotein vaaE [imported] - Mycoplasma pulmonis (strain UAB CTIP) (fragment)
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: P90577
R:Chambaud, I., Heilig, R., Ferris, S., Barbe, V., Samson, D., Galissou, F., Moszer, I.,
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis
A:Reference number: A9512; MUID:21267165; PMID:11353084
A:Accession: P90577
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-314 <KIR>
A:Cross-references: UNIPROT:O98043; GB:AL445566; PID:g14089940; PIDN:CAC1369.1; GSPDB:c:
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPu_5260
A:Genetic code: SGC3

```

```

Oy      2 VKEPPOQKYQRLH-----VQELTTEVEKIKITTVKESATBEKLTPLV 44
        ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      246 IFEQKAKFKFDVIQDARTKLDQLTTKLEKIKS--EKENIEKLDPII 290

```

RESULT 14  
JC5368  
dynactin 1 - mouse  
N:Alternate names: p150 Glued  
C:Species: Mus musculus (house mouse)  
C:Date: 28-May-1997 #sequence\_revision 18-Jul-1997 #text\_change 09-Jul-2004  
C:Accession: J05368  
R:Jiang, W.; Weber, J.S.; Tokito, M.K.; Holzbaur, E.L.F.; Weisler, M.H.  
Biochem. Biophys. Res. Commun. 231, 344-347, 1997  
A:Title: Mouse p150Glued (dynactin 1) cDNA sequence and evaluation as a candidate for th  
A:Reference number: J05368; MUID:97223454; PMID:9070275  
A:Accession: J05368  
A:Molecule type: mRNA  
A:Residues: 1-1281 <UN>  
A:Cross-references: UNIPROT:O08788; GB:U060312; NID:g2104494; PID:AMB57773.1; PID:g21044  
A:Experimental source: brain  
C:Comment: This protein is a member of the oligomeric dynactin complex that is required  
C:Genetics:  
A:Gene: Dctn1  
A:Map position: 6

```

QY      4 ETPOQKYORLLHEVQELTTEVEKIKTTVKESATEEKLTPVLLAKOL 49
      | : | : | | | | : | : : | | |
DB     319 ESLQGEVATLKERVDELTTDLILKAEIEKSGSDGAASSYOL-KOL 363

```

RESULT 15  
T32222  
hypothetical protein T33B12.7 - *Caenorhabditis elegans*  
CISpecies: *Caenorhabditis elegans*

C|Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C|Accession: T33222  
R|Davidson, S.; Wohlmann, P.; Giliam, B.  
submitted to the EMBL Data Library, September 1997  
A|Description: The sequence of C. elegans cosmid T33B12.  
A|Reference number: Z21137  
A|Accession: T332222  
A|Status: preliminary; translated from GB/EMBL/DBJ  
A|Residues: 1-296 <DAV>  
A|Molecule type: DNA  
A|Cross-references: UNIRROT:O17002; EMBL:AF022982; P11N:AB6938.1; GSPDB:GN00023; CESP  
A|Experimental source: strain Bristol N2; clone T33B12  
C|Genetics:  
A|Gene: CESP:T33B12.7  
A|Map position: 5  
A|Introns: 67/3

```
QY      3 KETPQCKQRLNHEVQELTTEVEK- IKTIVKESAT-----EKLTPVLLAK 47
      |||::: ||::: ||: |||
Db     132 KKTDPQNFK--EIEAIRQLKEKVNNEVQKATLTKTEREKHQKSEKLTPTLLK 186
```

Search completed: November 3, 2005, 22:04:16  
Job time : 18.9091 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 3, 2005, 21:51:39 ; Search time 15.4011 Seconds

(without alignments)  
145.410 Million cell updates/sec

Title: 09782816-1-22

Perfect score: 138

Sequence: 1 EVKIKTTVKESATEEKLTPVLLAKQIAL 30

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfiletest.pep.\*Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	37.0	136	4	US-09-270-767-36154
2	51	37.0	136	4	US-09-270-767-51371
3	50	36.2	938	4	US-09-637-145-2
4	49.5	35.9	119	4	US-09-621-976-4332
5	49	35.5	413	4	US-09-107-532A-6433
6	49	35.5	1027	4	US-09-107-532A-6675
7	48.5	35.1	124	4	US-09-513-999C-6026
8	48.5	35.1	280	3	US-09-470-512A-14
9	48.5	35.1	451	4	US-09-270-767-45139
10	48	34.8	1200	4	US-09-107-532A-6433
11	48	34.8	1428	4	US-09-644-827B-8
12	48	34.8	167	4	US-09-644-827B-7
13	47.5	34.4	314	4	US-09-166-350-13
14	47.5	34.4	314	4	US-09-902-540-16180
15	47.5	34.4	403	4	US-09-976-594-421
16	47	34.1	103	4	US-09-732-210-1198
17	47	34.1	180	4	US-09-648-004-2
18	47	34.1	180	4	US-10-272-419-2
19	47	34.1	289	4	US-09-071-035-480
20	47	34.1	317	3	US-09-134-001C-3712
21	47	34.1	387	4	US-09-248-796A-1845
22	47	34.1	424	4	US-09-286-981B-14
23	47	34.1	425	4	US-09-286-981B-13
24	47	34.1	426	4	US-09-286-981B-12
25	47	34.1	450	4	US-09-071-035-256
26	47	34.1	708	4	US-09-134-000C-5994
27	47	34.1	894	4	US-09-071-035-248

28	47	34.1	940	4	US-09-328-352-8165	Sequence 8165, Ap
29	47	34.1	962	4	US-09-071-035-246	Sequence 246, App
30	47	34.1	962	4	US-09-071-035-250	Sequence 250, App
31	47	34.1	962	4	US-09-071-035-254	Sequence 254, App
32	47	34.1	962	4	US-09-071-035-470	Sequence 470, App
33	47	34.1	962	4	US-09-071-035-474	Sequence 474, App
34	47	34.1	962	4	US-09-071-035-478	Sequence 478, App
35	47	34.1	970	4	US-09-134-000C-5691	Sequence 5691, Ap
36	46.5	33.7	1964	3	US-09-467-997-1	Sequence 1, Appl1
37	46	33.3	192	4	US-09-248-796A-20678	Sequence 20678, A
38	46	33.3	366	4	US-09-134-000C-4956	Sequence 4956, Ap
39	46	33.3	757	3	US-08-434-000A-6	Sequence 6, Appl1
40	46	33.3	757	3	US-09-312-157-6	Sequence 6, Appl1
41	46	33.3	757	3	US-09-717-888-6	Sequence 6, Appl1
42	46	33.3	1896	4	US-09-949-016-9508	Sequence 9508, Ap
43	46	32.6	85	4	US-09-732-210-1119	Sequence 1119, Ap
44	45	32.6	103	4	US-09-286-981B-1	Sequence 1, Appl1
45	45	32.6	114	4	US-09-286-981B-19	Sequence 19, Appl1

## ALIGNMENTS

```
RESULT 1
US-09-270-767-36154
; Sequence 36154, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 36154
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-51371

Query Match      37.0%; Score 51; DB 4; Length 136;
Best Local Similarity 40.0%; Pred. No. 7.6;
Matches 14; Conservative 6; Mismatches 9; Indels 6; Gaps 1;

Cy      2 VEKIKTT-----VKESATEEKLTPVLLAKQIAL 30
Db      37 IDKIKTTPTANPKTKRPTSNMKLSAVLLATALLAL 71

RESULT 2
US-09-270-767-51371
; Sequence 51371, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 51371
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-51371

Query Match      37.0%; Score 51; DB 4; Length 136;
Best Local Similarity 40.0%; Pred. No. 7.6;
Matches 14; Conservative 6; Mismatches 9; Indels 6; Gaps 1;
```

Db 37 IDKVTPTANPKPTPTNNMKLSAVLIALIAL 71

RESULT 3  
US-09-637-145-2  
; Sequence 2, Application US/09637145  
; Patent No. 6673587  
; GENERAL INFORMATION:  
; APPLICANT: EVANS, RONALD M.  
; APPLICANT: KAO, HUNG-YING  
; APPLICANT: DOMNES, MICHAEL  
; APPLICANT: ORDENTLICH, PETER  
; TITLE OF INVENTION: NOVEL HISTONE DEACETYLASE, AND USES THEREFOR  
; FILE REFERENCE: SAL3000  
; CURRENT APPLICATION NUMBER: US/09/637,145  
; CURRENT FILING DATE: 2000-08-11  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 938  
; TYPE: PRT  
; ORGANISM: Mus sp.  
US-09-637-145-2

Query Match 36.2%; Score 50; DB 4; Length 938;  
Best Local Similarity 44.0%; Pred. No. 94;  
Matches 11; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

OY 6 KTTYKESATEKLTPTVLAKQAL 30  
Db 100 RSAVASSVVKOKLAELVILKKQOAL 124

RESULT 4  
US-09-621-976-4332  
; Sequence 4332, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 4332  
; LENGTH: 119  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 66  
; OTHER INFORMATION: Xaa = \*, Trp  
US-09-621-976-4332

Query Match 35.9%; Score 49.5; DB 4; Length 119;  
Best Local Similarity 40.6%; Pred. No. 11;  
Matches 13; Conservative 7; Mismatches 9; Indels 3; Gaps 1;

OY 1 EVEKIKTVKESATEKLTPTVLAKQ--LAA 29  
Db 62 ELDKKTITMKSSVQECVSTISSKDEPPLAA 93

RESULT 5  
US-09-107-532A-6433  
; Sequence 6433, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS: GENOME THERAPEUTICS CORPORATION  
; ADDRESS: 100 Beaver Street  
; City: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD/ROM ISO9660  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arinello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 6433:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 413 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (B) LOCATION 1...413  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6433:  
US-09-107-532A-6433  
Query Match 35.5%; Score 49; DB 4; Length 413;  
Best Local Similarity 44.4%; Pred. No. 51;  
Matches 12; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

OY 1 EVEKIKTVKESATEKLTPT--VLAK 25  
Db 64 EERSQTAVALTKENRLTRGGVLT 90

RESULT 6  
US-09-107-532A-6675  
; Sequence 6675, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; City: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD/ROM ISO9660  
; COMPUTER: PC



```
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/107,532A
  FILING DATE: 30-Jun-1998
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 60/085,598
    FILING DATE: 14 May 1998
    APPLICATION NUMBER: 60/051571
    FILING DATE: July 2, 1997
  ATTORNEY/AGENT INFORMATION:
    NAME: Ariniello, Pamela Deneke
    REGISTRATION NUMBER: 40,489
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (781)893-5007
    TELEFAX: (781)893-8277
  INFORMATION FOR SEQ ID NO: 6675:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 1027 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      HYPOTHEetical: YES
      ORIGINAL SOURCE:
        ORGANISM: Enterococcus faecium
        FEATURE:
          NAME/KEY: misc feature
          LOCATION: (B) LOCATION 1...1027
          SEQUENCE DESCRIPTION: SEQ ID NO: 6675:
US-09-107-532A-6675

Query Match          35.1%; Score 49; DB 4; Length 1027;
Best Local Similarity 44.4%; Pred. No. 1.4e+02;
Matches 12; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

Qy      1 EVEKIKTVKESATEEKLTP--VLIAK 25
Db      788 EIERSGTAVVELTKENRLTPGVVLT 814

RESULT 7
US-09-513-999C-6026
; Sequence 6026, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 6026
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 46
; OTHER INFORMATION: Xaa-Gln or Arg
US-09-513-999C-6026

Query Match          35.1%; Score 48.5; DB 4; Length 124;
Best Local Similarity 39.3%; Pred. No. 15;
Matches 11; Conservative 10; Mismatches 6; Indels 1; Gaps 1;

Qy      1 EVEKIKTVKESATEEKLTPVLIAKOL 27
```

```
Db      40 EEKVKKSVADAKKGQKQDVCIVLAKEM 67

RESULT 8
US-09-470-512A-14
; Sequence 14, Application US/09470512A
; Patent No. 6378652
; GENERAL INFORMATION:
; APPLICANT: PhageTech, Inc.
; TITLE OF INVENTION: Compositions and methods involving an essential staphylococcus Au
; FILE REFERENCE: 21715/1000
; CURRENT APPLICATION NUMBER: US/09/470,512A
; CURRENT FILING DATE: 1999-12-12
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 14
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-470-512A-14

Query Match          35.1%; Score 48.5; DB 3; Length 280;
Best Local Similarity 33.3%; Pred. No. 39;
Matches 11; Conservative 8; Mismatches 9; Indels 5; Gaps 1;

Qy      2 VEKIKTVKESATEEKL-----TPVLIAKQDA 29
Db      171 VREIKNSIQDTLEKLMVKTTPVLMDDIGA 203

RESULT 9
US-09-270-767-45139
; Sequence 45139, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homberger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 45139
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-45139

Query Match          35.1%; Score 48.5; DB 4; Length 451;
Best Local Similarity 39.3%; Pred. No. 67;
Matches 11; Conservative 11; Mismatches 5; Indels 1; Gaps 1;

Qy      1 EVEKIKTVKESATEEKLTPVLIAKOL 27
Db      79 EEKVKKSLKQAQKNDRTCVIIAKEI 106

RESULT 10
US-09-107-532A-6432
; Sequence 6432, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: GENOM THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
```

```

; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Denoche
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6432:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8) LOCATION 1...107
; SEQUENCE DESCRIPTION: SEQ ID NO: 6432:
US-09-107-532A-6432

Query Match 34.8%; Score 48; DB 4; Length 107;
Best Local Similarity 47.6%; Pred. No. 15;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 5 IKTVKESATEKLTVPVLAK 25
:|::|||:|
Db 84 LKNIVSKNIEIKITPVQDK 104

RESULT 11
US-09-644-827B-8
; Sequence 8, Application US/09644827B
; Patent No. 6762283
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: SCHUCHMANN, Marcus
; APPLICANT: GONCHAROV, Tanya
; TITLE OF INVENTION: Caspase-8 Interacting Proteins
; FILE REFERENCE: WALLACH=26
; CURRENT APPLICATION NUMBER: US/09/644,827B
; CURRENT FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 132105
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: 127721
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-644-827B-8

Query Match 34.8%; Score 48; DB 4; Length 1200;
```

```

; Best Local Similarity 42.3%; Pred. No. 2.4e+02;
; Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 5 IKTVKESATEKLTVPVLAKQAL 30
:|::|||:|
Db 148 VACAVASSVVKQLAEVILKKQNAL 173

RESULT 12
US-09-644-827B-7
; Sequence 7, Application US/09644827B
; Patent No. 6762283
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: SCHUCHMANN, Marcus
; APPLICANT: GONCHAROV, Tanya
; TITLE OF INVENTION: Caspase-8 Interacting Proteins
; FILE REFERENCE: WALLACH=26
; CURRENT APPLICATION NUMBER: US/09/644,827B
; CURRENT FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 132105
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: 127721
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-644-827B-7

Query Match 34.8%; Score 48; DB 4; Length 1428;
Best Local Similarity 42.3%; Pred. No. 2.9e+02;
Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 5 IKTVKESATEKLTVPVLAKQAL 30
:|::|||:|
Db 374 VACAVASSVVKQLAEVILKKQNAL 399

RESULT 13
US-09-166-350-13
; Sequence 13, Application US/09166350A
; Patent No. 6440663
; GENERAL INFORMATION:
; APPLICANT: Scamman, Matthew
; APPLICANT: Chen, Yao
; APPLICANT: Stoeckert, Elisabeth
; APPLICANT: Old, Lloyd
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alex
; TITLE OF INVENTION: Renal Cancer Associated Antigens and
; FILE REFERENCE: L0461/7051
; CURRENT APPLICATION NUMBER: US/09/166,350A
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: US 09/166,350
; EARLIER FILING DATE: 1998-10-05
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 167
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-166-350-13

Query Match 34.4%; Score 47.5; DB 4; Length 167;
Best Local Similarity 40.6%; Pred. No. 30;
Matches 13; Conservative 7; Mismatches 9; Indels 3; Gaps 1;

Qy 1 EKKIKTVKESATEKLTVPVLAKQ--LAA 29
:|::|||:|
```

Db 100 ELDKWTMKSVOECVSTISSKEDPLAA 131

RESULT 14

US-09-902-540-16180

; Sequence 16180, Application US/09902540

; Patent No. 6833447

; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15869)B

; CURRENT APPLICATION NUMBER: US/09/902,540

; PRIOR FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883

; NUMBER OF SEQ ID NOS: 16825

; SEQ ID NO 16180

; LENGTH: 314

; TYPE: PRT

; ORGANISM: Myxococcus xanthus

US-09-902-540-16180

Query Match 34.4%; Score 47.5; DB 4; Length 314;

Best Local Similarity 50.0%; Pred. No. 61;

Matches 13; Conservative 3; Mismatches 9; Indels 1; Gaps 1;

3 EKITTVESATEEKLTP-VILAKQL 27

Db 244 EALKTFVEKATKRVTSVVLNPL 269

RESULT 15

US-09-976-594-421

; Sequence 421, Application US/09976594

; Patent No. 6673549

; GENERAL INFORMATION:

; APPLICANT: Furness, Michael

; APPLICANT: Buchbinder, Jenny

; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS

; FILE REFERENCE: PA-0041 US

; CURRENT APPLICATION NUMBER: US/09/976,594

; PRIOR FILING DATE: 2001-10-12

; PRIOR APPLICATION NUMBER: 60/240,409

; NUMBER OF SEQ ID NOS: 1143

; SOFTWARE: PERL Program

; SEQ ID NO 421

; LENGTH: 403

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. 6673549 1756951CD1

US-09-976-594-421

Query Match 34.4%; Score 47.5; DB 4; Length 403;

Best Local Similarity 40.6%; Pred. No. 81;

Matches 13; Conservative 7; Mismatches 9; Indels 3; Gaps 1;

1 EKEKITTVESATEEKLTPVILAKQ--LAA 29

Db 62 ELDKWTMKSVOECVSTISSKEDPLAA 93

Search completed: November 3, 2005, 22:05:55  
Job time : 16.4011 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: November 3, 2005, 21:53:15, Search time 54.0642 Seconds  
(without alignments)  
232.174 Million cell updates/sec

Title: 09782816-1-22

Perfect score: 138  
Sequence: 1 EWEKIKTVKESATEEKLTPVLAKQALAL 30

Scoring table:  
BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database:

Published Applications AA:\*

- 1: /cgn2\_6/prodata/1/pubppaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/prodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/prodata/1/pubppaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/prodata/1/pubppaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/prodata/1/pubppaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/prodata/1/pubppaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/prodata/1/pubppaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/prodata/1/pubppaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/prodata/1/pubppaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/prodata/1/pubppaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/prodata/1/pubppaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/prodata/1/pubppaa/US09D\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/prodata/1/pubppaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/prodata/1/pubppaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/prodata/1/pubppaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/prodata/1/pubppaa/US10D\_PUBCOMB.pep:\*
- 17: /cgn2\_6/prodata/1/pubppaa/US10E\_PUBCOMB.pep:\*
- 18: /cgn2\_6/prodata/1/pubppaa/US10\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/prodata/1/pubppaa/US11A\_PUBCOMB.pep:\*
- 20: /cgn2\_6/prodata/1/pubppaa/US11\_NEW\_PUB.pep:\*
- 21: /cgn2\_6/prodata/1/pubppaa/US60\_NEW\_PUB.pep:\*
- 22: /cgn2\_6/prodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	138	100.0	134	14	US-10-106-698-6730 Sequence 6730, Ap
2	138	100.0	401	16	US-10-408-765A-1369 Sequence 1369, Ap
3	138	100.0	406	10	US-09-782-816A-53 Sequence 53, Appl
4	138	100.0	465	9	US-09-925-298-676 Sequence 676, App
5	138	100.0	465	14	US-10-102-806-676 Sequence 676, App
6	135	97.8	183	10	US-09-782-816A-54 Sequence 54, Appl
7	135	97.8	224	16	US-10-425-115-31831 Sequence 31831, Appl
8	133	96.4	52	10	US-09-782-816A-51 Sequence 51, Appl
9	103	74.6	23	10	US-09-782-816A-1 Sequence 1, Appl
10	56	40.6	334	18	US-10-501-282-4734 Sequence 4734, Ap
11	51	37.0	447	15	US-10-335-977-7239 Sequence 7239, Ap

12	51	37.0	773	15	US-10-335-977-7240 Sequence 7240, Ap
13	51	37.0	932	15	US-10-282-122A-47253 Sequence 47253, A
14	51	37.0	10917	17	US-10-732-923-20606 Sequence 20606, A
15	50	36.2	211	16	US-10-425-115-212374 Sequence 212374, A
16	50	36.2	640	15	US-10-282-122A-44691 Sequence 44691, A
17	50	36.2	773	9	US-09-815-242-11330 Sequence 11330, A
18	50	36.2	773	15	US-10-282-122A-58722 Sequence 58722, A
19	50	36.2	855	9	US-09-817-913-13 Sequence 13, Appl
20	50	36.2	855	14	US-10-172-094-9 Sequence 9, Appl
21	50	36.2	855	15	US-10-189-818B-4 Sequence 4, Appl
22	50	36.2	855	16	US-10-870-587-13 Sequence 13, Appl
23	50	36.2	902	16	US-10-437-963-105564 Sequence 105564, A
24	50	36.2	912	10	US-09-800-187-12 Sequence 12, Appl
25	50	36.2	915	13	US-10-175-559-2 Sequence 2, Appl
26	50	36.2	915	14	US-10-087-192-1893 Sequence 1893, Ap
27	50	36.2	952	15	US-10-087-192-1893 Sequence 7, Appl
28	50	36.2	952	15	US-10-360-534-7 Sequence 1176, Ap
29	50	36.2	1642	17	US-10-741-600-1176 Sequence 1176, Ap
30	50	36.2	1642	17	US-10-741-600-1178 Sequence 1178, Ap
31	50	36.2	3024	16	US-10-408-765A-1977 Sequence 1977, Ap
32	50	36.2	3396	16	US-10-788-792-1170 Sequence 1170, App
33	50	36.2	3396	17	US-10-741-600-1172 Sequence 1172, Ap
34	50	36.2	3396	17	US-10-741-600-1173 Sequence 1173, Ap
35	50	36.2	3396	18	US-10-631-467-773 Sequence 773, App
36	49	35.5	87	9	US-09-864-761-43703 Sequence 43703, A
37	49	35.5	279	15	US-10-282-122A-57857 Sequence 57857, A
38	49	35.5	291	16	US-10-767-701-40287 Sequence 40287, A
39	49	35.5	307	11	US-09-972-211-73 Sequence 73, Appl
40	49	35.5	307	15	US-10-096-625-73 Sequence 73, Appl
41	49	35.5	340	11	US-09-972-211-72 Sequence 72, Appl
42	49	35.5	340	15	US-10-096-625-72 Sequence 72, Appl
43	49	35.5	340	16	US-10-723-860-1930 Sequence 1930, Ap
44	49	35.5	340	18	US-10-756-149-5208 Sequence 5208, Ap
45	49	35.5	352	18	US-10-450-763-52703 Sequence 52703, A

## ALIGNMENTS

RESULT 1  
US-10-106-698-6730  
Sequence 6730, Application US/10106698  
Publication No. US20030109690A1  
GENERAL INFORMATION:  
APPLICANT: Ruben et al.  
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide  
FILE REFERENCE: P4005P1  
CURRENT APPLICATION NUMBER: US/10/106,698  
PRIOR FILING DATE: 2002-03-27  
PRIOR APPLICATION NUMBER: PCT/US00/26524  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 60/157,137  
PRIOR FILING DATE: 1999-09-29  
PRIOR APPLICATION NUMBER: US 60/163,280  
PRIOR FILING DATE: 1999-11-03  
NUMBER OF SEQ ID NOS: 8564  
SOFTWARE: PatentIn Ver. 3.0  
SEQ ID NO 6730  
LENGTH: 134  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (126)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-106-698-6730

Query Match 100.0%, Score 138, DB 14, Length 134,  
Best Local Similarity 100.0%, Pred. No. 2.3e-11,  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 EWEKIKTVKESATEEKLTPVLAKQALAL 30  
|||||

```
Db      19  EVEKITTVKESATEEKLTPVLAQKQAL 48

RESULT 2
US-10-408-765A-1369
; Sequence 1369, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Boin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Marnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1369
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1369

Query Match      100.0%; Score 138; DB 16; Length 401;
Best Local Similarity 100.0%; Pred. No. 8.4e-11;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1  EVEKITTVKESATEEKLTPVLAQKQAL 30
Db      116 EVEKITTVKESATEEKLTPVLAQKQAL 145

RESULT 3
US-09-782-816A-53
; Sequence 53, Application US/09782816A
; Publication No. US20030032771A1
; GENERAL INFORMATION:
; APPLICANT: Sharp, David J.
; APPLICANT: Rogers, Gregory C.
; APPLICANT: Scholey, Jonathan M.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
; FILE REFERENCE: UC069.001A
; CURRENT APPLICATION NUMBER: US/09/782,816A
; CURRENT FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-816A-53

Query Match      100.0%; Score 138; DB 10; Length 406;
Best Local Similarity 100.0%; Pred. No. 8.5e-11;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1  EVEKITTVKESATEEKLTPVLAQKQAL 30
Db      121 EVEKITTVKESATEEKLTPVLAQKQAL 150

RESULT 4
US-09-925-298-676
; Sequence 676, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

Db      19  EVEKITTVKESATEEKLTPVLAQKQAL 48

TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 676
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
FEATURE:
; NAME/KEY: SITE
; LOCATION: (5)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (16)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-676

Query Match      100.0%; Score 138; DB 9; Length 465;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1  EVEKITTVKESATEEKLTPVLAQKQAL 30
Db      160 EVEKITTVKESATEEKLTPVLAQKQAL 209

RESULT 5
US-10-102-806-676
; Sequence 676, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 676
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
FEATURE:
; NAME/KEY: SITE
; LOCATION: (5)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (16)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-676

Query Match      100.0%; Score 138; DB 14; Length 465;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 EVEKIKITVAKESATEEKLTPVLLAKQLAAL    30
         |||||
Db       180 EVEKIKITVAKESATEEKLTPVLLAKQLAAL    209
```

```

RESULT 6
US-09-782-816A-54
; Sequence 54, Application US/09782816A
; Publication No. US20030032771A1
; GENERAL INFORMATION:
; APPLICANT: Sharp, David J.
; APPLICANT: Rogers, Gregory C.
; APPLICANT: Schooley, Jonathan M.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
; TITLE OF INVENTION: PROLIFERATION
; FILE REFERENCE: UC069, 001A
; CURRENT APPLICATION NUMBER: US/09/782, 816A
; CURRENT FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54
; LENGTH: 183
; TYPE: PRP
; ORGANISM: Mus musculus
US-09-782-816A-54

```

Query Match	97.8%	Score 135;	DB 10;	Length 183;
Best Local Similarity	96.7%	Pred. No. 8.7e-11;		
Matches 29; Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0;

```
QY      1 EVEKIKITVYESATEEKLTPVLAKOQLAAL  30  
        |||||  
Db      116 EVEKIKITVYESATEEKLTPVLAKOQLAAL 145
```

```

RESULT 7
US-10-425-115-315831
; Sequence 315831, Application US/10425115
; Publication NO. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 315831
; LENGTH: 224
; TYPE: prt
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_51109C.1.pep
; US-10-425-115-315831

```

Query Match	97.8%	Score 135;	DB 16;	Length 224;
Best Local Similarity	96.7%	Pred. No. 1.1e-10;		
Matches 29; Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0;

```
QY      1  EWEKIKTTVESATEEKLTPVLLAKQIAL 30
        |||||:|||||
Db      69  EWEKIKTTVESATEEKLTPVLLAKQIAL 98
```

RESULT 8  
US-09-782-816A-51  
; Sequence 51, Application US/09782816A  
; Publication No. US20030032771A1  
; GENERAL INFORMATION:  
; APPLICANT: Sharp, David J.

```

1  APPLICANT:  Rogers, Gregory C.
2  APPLICANT:  Scholey, Jonathan M.
3  TITLE OF INVENTION:  PEPTIDE INHIBITORS OF CELLULAR
4  TITLE OF INVENTION:  PROLIFERATION
5  FILE REFERENCE:  UC069, 001A
6  CURRENT APPLICATION NUMBER:  US/09/782,816A
7  CURRENT FILING DATE:  2001-02-14
8  NUMBER OF SEQ. ID NOS:  56
9  SOFTWARE:  FastSeq for Windows Version 4.0
10 SEQ ID NO 51
11 LENGTH:  52
12 TYPE:  PRT
13 ORGANISM:  Unknown
14 FEATURE:
15 NAME/KEY:  UNSURE
16 LOCATION:  44
17 OTHER INFORMATION:  Xaa = Val or Leu
18 OTHER INFORMATION:  A sequence conserved among Homo sapiens and Mus
19 OTHER INFORMATION:  musculus.
20 US-09-782-816A-51

```

Query Match	96.4%	Score 133;	DB 10;	Length 52;
Best Local Similarity	96.7%	Pred. No. 3.8e-11;		
Matches 29; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

```
QY      1 EVEKIKTTVESATEEKLTPVLAKQIAL 30
        |||||
Db      23 EVEKIKTTVESATEEKLTPVXLAKQIAL 52
```

```

RESULT 9
US-09-782-816A-1
: Sequence 1, Application US/09782816A
: Publication No. US20030032771A1
: GENERAL INFORMATION:
: APPLICANT: Sharp, David J.
: APPLICANT: Rogers, Gregory C.
: APPLICANT: Scholey, Jonathan M.
: TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
: TITLE OF INVENTION: PROLIFERATION
: FILE REFERENCE: UC069.001A
: CURRENT APPLICATION NUMBER: US/09/782, 816A
: CURRENT FILING DATE: 2001-02-14
: NUMBER OF SEQ. ID NOS: 56
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 23
: TYPE: PRT
: ORGANISM: Unknown
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: 22
: OTHER INFORMATION: Xaa = Val or Leu
: OTHER INFORMATION: The sequence is a Homo sapiens sequence when Xaa
: OTHER INFORMATION: represents Leu and a Mus musculus sequence when.
US-09-782-816A-1

```

Query Match	74.6%	Score 103;	DB 10;	Length 23;
Best Local Similarity	95.7%	Pred. No. 2.3e-07;		
Matches 22;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

```
QY      1 EVEKIKITVKESATEEKLTPVL 23
        |||||
Db      1 EVEKIKITVKESATEEKLTPVL 23
```

RESULT 10  
US-10-501-282-4734  
; Sequence 4734, Application US/10501282;  
; Publication No. US20050203280A1  
; GENERAL INFORMATION:  
; APPLICANT: MCMICHAEL, JOHN CALHOUN

```

1  APPLICANT: ZAGURSKY, ROBERT JOHN
2  APPLICANT: RUSSELL, DAVID PARRISH
3  APPLICANT: FLETCHER, LEAH DIANE
4  TITLE OF INVENTION: ALLOTOCOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
5  TITLE OF INVENTION: POLYPEPTIDE ANTIGENS, IMMUNOGENIC COMPOSITIONS AND USES THEREOF
6  FILE REFERENCE: AM100780 L2
7  CURRENT APPLICATION NUMBER: US/10/501,282
8  CURRENT FILING DATE: 2004-07-09
9  PRIOR APPLICATION NUMBER: 60/333,777
10 PRIOR FILING DATE: 2001-11-29
11 PRIOR APPLICATION NUMBER: 60/426,742
12 PRIOR FILING DATE: 2002-11-18
13 PRIOR APPLICATION NUMBER: PCT/US02/36123
14 PRIOR FILING DATE: 2002-11-25
15 NUMBER OF SEQ ID NOS: 6653
16 SOFTWARE: Patentin version 3.2
17 SEQ ID NO 4734
18 LENGTH: 334
19 TYPE: prt
20 ORGANISM: Alloiococcus otitidis
21 US-10-501-282-4734

```

		Query Match	40.6%	Score 56,	DB 18;	Length 334;
		Best Local Similarity	34.5%;	Pred. No. 21;		
	Matches	10; Conservative	11; Mismatches	8; Indels	0; Gaps	0.
Oy		2 VKETKYKESATBEKLTPVLAKOTAL	30			
	:	:::: :	: :	:	:	:
Dz	244	IDQIKSTIKSGSVKTKLGLALKOLSSL	272			

```

1      RESULT 11
2      US-10-335-977-7239
3      Sequence 7239: Application US/10335977
4      Publication No. US20040052799A1
5      GENERAL INFORMATION:
6      APPLICANT: DOUGLAS SMITH et al
7      TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
8      RELATING TO HELICOBACTER PYLORI FOR
9      DIAGNOSTICS AND THERAPEUTICS
10     NUMBER OF SEQUENCES: 10031
11     CORRESPONDENCE ADDRESS:
12     ADDRESSEE: LAHIVE & COCKFIELD
13     STREET: 28 State Street
14     CITY: Boston
15     STATE: Massachusetts
16     COUNTRY: USA
17     ZIP: 02109-1875
18     COMPUTER READABLE FORM:
19     MEDIUM TYPE: CD-ROM ISO9660
20     COMPUTER: IBM PC Compatible
21     OPERATING SYSTEM: Windows NT 4.0
22     SOFTWARE: UNIX
23     CURRENT APPLICATION DATA:
24     APPLICATION NUMBER: US/10/335,977
25     FILING DATE: 30-Dec-2002
26     PRIOR APPLICATION DATA:
27     APPLICATION NUMBER: 08/993,002
28     FILING DATE: 17-Dec-1997
29     ATTORNEY/AGENT INFORMATION:
30     NAME: Mandagouras, Amy E.
31     REGISTRATION NUMBER: 36,207
32     REFERENCE/DOCKET NUMBER: GTN-018
33     TELECOMMUNICATION INFORMATION:
34     TELEPHONE: (617) 227-7400
35     TELEFAX: (617) 742-4214
36     INFORMATION FOR SEQ ID NO: 7239:
37     SEQUENCE CHARACTERISTICS:
38     LENGTH: 447 amino acids
39     TYPE: amino acid
40     TOPOLOGY: linear
41     MOLECULE TYPE: protein
42     HYPOTHETICAL: YES

```

```

? ORIGINAL SOURCE:
? ORGANISM: Helicobacter pylori
?
? FEATURE:
? NAME/KEY: misc feature
? LOCATION: (8) LOCATION 1...447
? SEQUENCE DESCRIPTION: SEQ ID NO: 7239
US-10-335-977-7239
```

Query Match	37.0%	Score 51	DB 15	Length 447
Best Local Similarity	43.5%	Pred. No.	1.5e+02	
Matches 10	Conservative 6	Mismatches 7	Indels 0	Gaps 0

```
QY      5 IKTVKESATEEKLTPVLLAKOL 27  
       :|::||:||::|:  
Db     202 LKQFVKDSAKKELTPIIAFKSM 224
```

```

      RESULT 12
US-10-335-977-7240
; Sequence 7240, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 7240:
SEQUENCE CHARACTERISTICS:
LENGTH: 773 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEetical: YES
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...773
SEQUENCE DESCRIPTION: SEQ ID NO: 7240:
US-10-335-977-7240
Query Match          37.0%; Score 51; DB 15; Length 773;
Best Local Similarity 43.5%; Pred. No. 2.8e+02;
Matches    10; Conservative   6; Mismatches     7; Indels       0; Gaps       0.

```

```

APPLICANT: DOUGLAS SMITH et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO HELICOBACTER PYLORI FOR
DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 10031
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 7240:
SEQUENCE CHARACTERISTICS:
LENGTH: 773 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...773
SEQUENCE DESCRIPTION: SEQ ID NO: 7240:
US-10-335-977-7240
Query Match 37.0%; Score 51; DB 15; Length 773;
Best Local Similarity 43.5%; Pred.No. 2.8e+02;
Matches 10; Conservative 6; Mismatches 7; Indels 0; Gaps 0.

```





**This Page Blank (uspto)**



CC polypeptides; expression vectors and host cells comprising a TAR nucleic acid; an antibody specific for a TAR polypeptide; a peptide or organic molecule which binds to a TAR polypeptide; fusion proteins comprising a TAR polypeptide; and methods and compositions for the treatment or diagnosis of cancer in mammals. TAR polypeptides, nucleic acids, antibodies, antagonists, binding molecules and compositions are useful for diagnosing or treating a cell proliferative disorder associated with increased TAR expression, particularly cancers such as breast cancer, colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder cancer, pancreatic cancer, cervical cancer, cancers of the central nervous system, melanoma and leukaemia. TAR nucleic acids may further be used as hybridisation probes, in chromosome and gene mapping, in chromosome identification and in gene therapy. The present sequence represents a TAR polypeptide of the invention

SO Sequence 314 AA;

Query Match 100.0%; Score 251; DB 8; Length 314;  
Best Local Similarity 100.0%; Pred. No. 6e-21;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC 1 GVKEPPOOKYORLHVEQELTTEVEKIKTKVKSATREKTPVLLAKQALAL 52  
DB 7 GVKEPPOOKYORLHVEQELTTEVEKIKTKVKSATREKTPVLLAKQALAL 58

RESULT 2  
ABP98851  
ID ABP98851 standard; protein; 378 AA.  
AC ABP98851;  
XX  
XX 15-JUN-2003 (first entry)  
DE Human structural and cytoskeletal associated protein #42.  
XX  
XX Cytostatic; antiarteriosclerotic; anticonvulsant; nootropic; antianginal;  
KW neuroprotective; cerebroprotective; hypotensive; cardiant; osteoparitic;  
KW antiinflammatory; antiarthritic; vincidine; gene therapy; human; stroke;  
KW structural and cytoskeleton-associated protein; SCAP; cancer; angina;  
KW atherosclerosis; epilepsy; Huntington's disease; hypertension;  
KW heart failure; osteoporosis; osteoarthritis.  
XX  
XX Homo sapiens.  
OS  
PN WO2003031940-A2.  
XX  
PD 17-APR-2003.  
XX  
PF 10-OCT-2002; 2002WO-US032851.  
XX  
PR 12-OCT-2001; 2001US-0328931P.  
PR 19-OCT-2001; 2001US-0360681P.  
PR 02-NOV-2001; 2001US-0343896P.  
PR 09-NOV-2001; 2001US-0346308P.  
PR 16-NOV-2001; 2001US-0332385P.  
PR 07-DEC-2001; 2001US-0340776P.  
PR 11-JAN-2002; 2002US-0347703P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
PA  
PI Becha SD, Blattia V, Blake JD, Borowsky ML, Burrill JD, Chang H,  
PI Chawla NK, Elliott US, Emerling BM, Forsythe TJ, Gorrard AE,  
PI Griffin VA, Hafalia AD, Ho A, Ison CH, Kable AE, Khare R, Lal PG,  
PI Lee S, Lee EA, Lee SY, Lehr-Mason PM, Li JX, Lindquist EA, Luo W,  
PI Marquis JB, Ramkumar J, Richardson TW, Sprague WW, Swarnakar A,  
PI Tang YT, Warren BA, Yang J, Yue H, Zebardjian Y, Zheng W;  
XX  
XX WPI; 2003-403125/38.  
DR N-PSDB; ACC44338.  
XX  
PT New human structural and cytoskeleton-associated proteins (SCAP) useful  
for diagnosing, treating and preventing diseases or conditions associated

PT with the aberrant SCAP expression e.g. cancer, osteoporosis, or epilepsy.  
XX  
XX Claim 1; Page 298; 361pp; English.  
XX  
XX This sequence represents a novel isolated human structural and  
CC cytoskeleton-associated protein (SCAP) polypeptide. The polypeptides and  
CC polynucleotides encoding them are useful in diagnosing, treating and  
CC preventing diseases or conditions associated with the decreased  
CC expression or over expression of SCAP, such as cell proliferative (e.g.  
CC cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's  
CC disease, stroke), heart (e.g. hypertension, heart failure, angina) and  
CC skeletal muscle disorders (e.g. osteoporosis, osteoarthritis) or viral  
CC infections. These are also useful in assessing the effects of exogenous  
CC compounds on the expression of nucleic acid and amino acid sequences of  
CC SCAP. The SCAP or its fragments are useful in screening compounds for  
CC effectiveness as agonist or antagonist of the polypeptides, or in  
CC altering the expression of the target polynucleotide and compounds that  
CC specifically bind to or modulate the activity of the polypeptide. The  
CC microarray is useful in monitoring or measuring protein-protein  
CC interactions, drug-target interactions, and gene expression profiles  
XX

SO Sequence 378 AA;

Query Match 100.0%; Score 251; DB 6; Length 378;  
Best Local Similarity 100.0%; Pred. No. 7.4e-21;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC 1 GVKEPPOOKYORLHVEQELTTEVEKIKTKVKSATREKTPVLLAKQALAL 52  
DB 71 GVKEPPOOKYORLHVEQELTTEVEKIKTKVKSATREKTPVLLAKQALAL 122

RESULT 3  
ADU6563  
ID ADU6563 standard; protein; 401 AA.  
XX  
XX ADU6563;  
AC  
XX 06-MAY-2004 (first entry)  
DE Human heat mitochondrial protein as a therapeutic target SegID1369.  
XX  
XX mitochondria; human; screening assay; diabetes mellitus;  
KW Huntington's disease; osteoarthritis;  
KW Leber's hereditary optic neuropathy; LHON;  
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;  
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;  
KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;  
KW osteopathic; ophthalmological; cytostatic.  
XX  
XX Homo sapiens.  
OS  
PN WO2003087768-A2.  
XX  
PD 23-OCT-2003.  
XX  
PF 04-APR-2003; 2003WO-US010870.  
XX  
PR 12-APR-2002; 2002US-0372843P.  
PR 17-JUN-2002; 2002US-0389967P.  
PR 20-SEP-2002; 2002US-0412418P.  
XX  
XX (MITO-) MITOKOR.  
PA  
PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM,  
PI Warnock DE;  
XX  
XX WPI; 2003-845369/78.  
DR  
XX  
PT Identifying a mitochondrial target for drug screening assays and for  
treating diseases associated with altered mitochondrial function,  
PT comprises detecting a modified polypeptide in a sample and correlating

PT with the disease.  
XX  
PS Claim 1; SEQ ID NO 1369; 180pp; English.  
XX  
CC This invention relates to novel mitochondrial targets that can be used  
CC for therapeutic intervention in treating a disease associated with  
CC altered mitochondrial function. Specifically, it refers to a method for  
CC identifying proteins of the human heart mitochondrial proteome that are  
CC useful for drug screening assays, as well as therapeutic targets. The  
CC present invention describes a method for identifying such proteins that  
CC can be used in the treatment of various diseases associated with altered  
CC mitochondrial function including diabetes mellitus, Huntington's disease,  
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial  
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy  
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these  
CC compositions have neuroprotective, neurotropic, antidiabetic,  
CC anticonvulsant, antirheumatic, osteopathic, ophthalmological and  
CC cytoprotective activities. This polypeptide sequence is a human heart  
CC mitochondrial protein of the invention.  
XX  
SQ Sequence 401 AA;  
  
Query Match 100.0%; Score 251; DB 7; Length 401;  
Best Local Similarity 100.0%; Pred. No. 7.9e-21;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GVKETPOOKYQRLHVEVQLTTEVEKIKTVKESATBEKLTPTVLLAKQLAAL 52  
DB 94 GVKETPOOKYQRLHVEVQLTTEVEKIKTVKESATBEKLTPTVLLAKQLAAL 145  
  
RESULT 4  
ABP53018  
ID ABP53018 standard; protein; 406 AA.  
XX  
AC ABP53018;  
XX  
DT 05-NOV-2002 (first entry)  
XX  
DE Human p50 amino acid sequence SEQ ID NO:53.  
XX  
XX Cellular proliferation inhibition; cytostatic; antiinflammatory; cancer;  
KM p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
KM sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
KM glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;  
KM glandular disorder; macropagal disorder; epithelial disorder;  
KM stromal disorder; blastocoeic disorder; angiogenic disorder;  
KM immunologic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200264779-A2.  
XX  
PD 22-AUG-2002.  
XX  
PP 21-JAN-2002; 2002WO-US001708.  
XX  
PR 14-FEB-2001; 2001US-00782816.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
PI Sharp DJ, Rogers GC, Scholey JM;  
XX  
DR WPI; 2002-657599/70.  
XX  
XX New peptide inhibitors of p50/dynaminin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.  
XX  
PS Disclosure; Fig 1; 55pp; English.  
XX  
CC The present invention describes an isolated peptide (I) comprising or

CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
CC terminal extensions. (I) have cytostatic and antiinflammatory activities  
CC and can be used as p50/dynaminin inhibitors and in gene therapy. The  
CC peptides, nucleic acid molecules and methods from the present invention  
CC are useful for treating cancer by inhibiting cellular proliferation, such  
CC as benign or malignant tumours (renal, liver, kidney, bladder, breast,  
CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
CC thyroid, hepatic carcinomas, sarcoma, glioblastomas, and various head and  
CC neck tumours); leukemias and lymphoid malignancies, other disorders such  
CC as neuronal, glial, astrocytal, hypothalamic and other glandular,  
CC macropagal, epithelial, stromal and blastocoeic disorders; and  
CC inflammatory, angiogenic and immunologic disorders. The present sequence  
CC represents human p50 which is given in the exemplification of the present  
CC invention  
XX  
SQ Sequence 406 AA;  
  
Query Match 100.0%; Score 251; DB 5; Length 406;  
Best Local Similarity 100.0%; Pred. No. 8.1e-21;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GVKETPOOKYQRLHVEVQLTTEVEKIKTVKESATBEKLTPTVLLAKQLAAL 52  
DB 99 GVKETPOOKYQRLHVEVQLTTEVEKIKTVKESATBEKLTPTVLLAKQLAAL 150  
  
RESULT 5  
AAB58968  
ID AAB58968 standard; protein; 465 AA.  
XX  
AC AAB58968;  
XX  
DT 27-MAR-2001 (first entry)  
XX  
DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 676.  
XX  
XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;  
KM neurotropic; neuroprotective; antiviral; antiallergic; hepatocytic;  
KM antidiabetic; antiinflammatory; antitumor; vulnery; anticonvulsant;  
KM antibacterial; antifungal; antiparasitic; cardiac; immune disorder;  
KM Addison's disease; allergy; autoimmune haemolytic anaemia;  
KM autoimmune thyroiditis; diabetes mellitus; Crohn's disease;  
KM multiple sclerosis; rheumatoid arthritis; ulcerative colitis;  
KM cardiovascular disorder; wound healing; neurological disease.  
XX  
OS Homo sapiens.  
XX  
FN WO200055173-A1.  
XX  
PD 21-SEP-2000.  
XX  
PP 08-MAR-2000; 2000WO-US005681.  
XX  
PR 12-MAR-1999; 99US-0124270P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Ruben SM;  
XX  
DR WPI; 2000-611515/58.  
XX  
DR N-PSDB; AAF21871.  
XX  
XX New human breast and ovarian cancer associated gene sequences and the  
PT polypeptides encoded by these genes, useful in the prevention, treatment  
PT and diagnosis of cancer, immune disorders, cardiovascular disorders and  
PT neurological diseases.  
XX  
PS Claim 11; Page 1126-1128; 1299pp; English.  
XX  
CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human  
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are  
CC associated with breast and ovarian cancer. Included in the invention are

CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the  
 CC isolation and characterisation of the DNA and protein sequences of the  
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist  
 CC or antagonist sequences exhibit cytostatic, immunosuppressive, neutrotropic,  
 CC neuroprotective; antiviral; antiallergic; hepatocytic; antidiabetic;  
 CC antifungal; antiparasitic and cardiant activity. The polynucleotide and  
 CC protein sequences are used in the diagnosis of cancer, particularly  
 CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists  
 CC and agonists may also be used in the diagnosis, prevention and treatment  
 CC of immune disorders e.g. Addison's disease, allergies, autoimmune  
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
 CC cardiovascular disorders such as myocardial ischaemia; wound healing;  
 CC neurological diseases such as cerebral anoxia and epilepsy; and  
 CC infectious diseases

CC Sequence 465 AA;

Query Match 100.0%; Score 251; DB 3; Length 465;  
 Best Local Similarity 100.0%; Pred. No. 9,4e-21;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVKEPPOQKYORLHVEQELTTEVEKIKTYVESATEEKLTPVLLAKQLAAL 52  
 |||||  
 Db 158 GVKEPPOQKYORLHVEQELTTEVEKIKTYVESATEEKLTPVLLAKQLAAL 209

RESULT 6  
 AAB53019  
 ID AAB53019 standard; protein, 183 AA.

AC AAB53019;

DT 05-NOV-2002 (first entry)

XX Mouse p50 amino acid sequence SEQ ID NO:54.

XX Cellular proliferation inhibitor; gene therapy; tumour; carcinoma;  
 KM p50 inhibitor; dynaminin inhibitor; lymphoid malignancy; neuronal disorder;  
 KM sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
 KM glioma; astrocytic disorder; hypochalamic disorder; inflammatory;  
 KM glandular disorder; macrophagal disorder; epithelial disorder;  
 KM stromal disorder; blastocoele disorder; angiogenic disorder;  
 KM immunologic disorder.

XX Mus musculus.

OS WO200264779-A2.

XX 22-AUG-2002.

XX 21-JAN-2002; 2002WO-US001708.

XX 14-FEB-2001; 2001US-00782816.

XX (REGC ) UNIV CALIFORNIA.

PI Sharp DJ, Rogers GC, Scholey JM;

XX WPI; 2002-657599/70.

XX New peptide inhibitors of p50/dynaminin useful for treating cancer by  
 PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
 PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
 PT immunologic disorders.

XX Disclosure; Fig 2; 55pp; English.

XX The present invention describes an isolated peptide (1) comprising or  
 CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
 CC sequences given in AAB52966 and AAB52967 and can have C-terminal and N-  
 CC terminal extensions. (1) have cytostatic and antiinflammatory activities

CC and can be used as p50/dynaminin inhibitors and in gene therapy. The  
 CC peptides, nucleic acid molecules and methods from the present invention  
 CC are useful for treating cancer by inhibiting cellular proliferation, such  
 CC as benign or malignant tumours (renal, liver, kidney, bladder, breast,  
 CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
 CC thyroid, hepatic carcinomas, sarcoma, glioblastomas, and various head and  
 CC neck tumours); leukemias and lymphoid malignancies, other disorders such  
 CC as neuronal, glial, astrocytic, hypochalamic and other glandular,  
 CC macrophagal, epithelial, stromal and blastocoele disorders; and  
 CC inflammatory, angiogenic and immunologic disorders. The present sequence  
 CC represents mouse p50 which is given in the exemplification of the present  
 CC invention

CC Sequence 183 AA;

Query Match 98.8%; Score 248; DB 5; Length 183;  
 Best Local Similarity 98.1%; Pred. No. 7,1e-21;  
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVKEPPOQKYORLHVEQELTTEVEKIKTYVESATEEKLTPVLLAKQLAAL 52  
 |||||  
 Db 94 GVKEPPOQKYORLHVEQELTTEVEKIKTYVESATEEKLTPVLLAKQLAAL 145

RESULT 7  
 AAB53016  
 ID AAB53016 standard; peptide, 52 AA.

AC AAB53016;

DT 05-NOV-2002 (first entry)

XX Cellular proliferation peptide inhibitor SEQ ID NO:51.

XX Cellular proliferation inhibitor; gene therapy; tumour; carcinoma;  
 KM p50 inhibitor; dynaminin inhibitor; lymphoid malignancy; neuronal disorder;  
 KM sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
 KM glioma; astrocytic disorder; hypochalamic disorder; inflammatory;  
 KM glandular disorder; macrophagal disorder; epithelial disorder;  
 KM stromal disorder; blastocoele disorder; angiogenic disorder;  
 KM immunologic disorder.

XX Homo sapiens.

OS Mus musculus.

OS Synthetic.

XX Key Location/Qualifiers  
 FH Misc-difference 44 /label= Leu, Val  
 FT /note= "Leu in humans and Val in Mus musculus"  
 FT  
 XX WO200264779-A2.

XX 22-AUG-2002.

XX 21-JAN-2002; 2002WO-US001708.

XX 14-FEB-2001; 2001US-00782816.

XX (REGC ) UNIV CALIFORNIA.

PI Sharp DJ, Rogers GC, Scholey JM;

XX WPI; 2002-657599/70.

XX New peptide inhibitors of p50/dynaminin useful for treating cancer by  
 PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
 PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
 PT immunologic disorders.

XX Claim 2; Page 31; 55pp; English.

XX The present invention describes an isolated peptide (1) comprising or

CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
CC terminal extensions. (1) have cytostatic and anti-inflammatory activities  
CC and can be used as p50/dynaminin inhibitors and in gene therapy. The  
CC peptides, nucleic acid molecules and methods from the present invention  
CC are useful for treating cancer by inhibiting cellular proliferation, such  
CC as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
CC neck tumors); leukemias and lymphoid malignancies, other disorders such  
CC as neuronal, glial, astrocytal, hypochalamic and other glandular,  
CC macrophagal, epithelial, stromal and blastocoelec disorders; and  
CC inflammatory, angiogenic and immunologic disorders. The present sequence  
CC represents a specifically claimed peptide inhibitor of cellular  
CC proliferation from the present invention

XX  
SQ Sequence 52 AA;

Query March 98.0%; Score 246; DB 5; Length 52;  
Best Local Similarity 98.1%; Pred. No. 2.8e-21;  
Matches 51; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVKETPOOKYQRLHVEVQLTTEVEKIKTYVESATEEKLTPVLLAKQALAL 52  
Db 1 GVKETPOOKYQRLHVEVQLTTEVEKIKTYVESATEEKLTPVLLAKQALAL 52

RESULT 8  
AAG75956  
ID AAG75956 standard; protein; 134 AA.

XX  
AC AAG75956;  
XX  
DT 03-SEP-2001 (first entry)

XX  
DE Human colon cancer antigen protein SEQ ID NO:6720.

XX  
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KW colorectal carcinoma; chromosome 12.

XX  
OS Homo sapiens.

XX  
PN WO200122920-A2.

XX  
PD 05-APR-2001.

XX  
PF 28-SEP-2000; 2000WO-US026524.

XX  
PR 29-SEP-1999; 99US-0157137P.  
PR 03-NOV-1999; 99US-0163280P.

XX  
PA (HUMA-) HUMAN GENOME SCI INC.

XX  
PI Ruben SM, Barash SC, Birse CE, Rosen CA;  
XX  
DR WPI; 2001-235357/24.  
DR N-PSDB; AAH35361.

XX  
PT Nucleic acid encoding 4277 human colon cancer-associated polypeptides,  
PT useful for preventing, diagnosing and/or treating colorectal cancers.

XX  
PS Claim 1; Page 8181; 9803pp; English.

XX  
AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the  
CC proteins are collectively known as colon cancer antigens. The colon  
CC cancer antigens have cytostatic activity and can be used in gene therapy  
CC and vaccine production. N and P may be used in the prevention, diagnosis  
CC and treatment of diseases associated with inappropriate P expression. For  
CC example, N and P may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of P by expressing inactive proteins or to  
CC supplement the patients own production of P. Additionally, N may be used

CC to produce the colon cancer-associated Ps, by inserting the nucleic acids  
CC into a host cell and culturing the cell to express the proteins. N and P  
CC can be used in the prevention, diagnosis and treatment of colorectal  
CC carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent  
CC sequences used in the exemplification of the present invention. N.B.  
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at  
CC time of publication, meaning no sequences are present for SEQ ID NO:1027  
CC to 1052, 7921 and 7922

XX  
SQ Sequence 134 AA;

Query March 92.0%; Score 231; DB 4; Length 134;  
Best Local Similarity 100.0%; Pred. No. 4.8e-19;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TPQOKYQRLHVEVQLTTEVEKIKTYVESATEEKLTPVLLAKQALAL 52  
Db 1 TPQOKYQRLHVEVQLTTEVEKIKTYVESATEEKLTPVLLAKQALAL 48

RESULT 9  
ABP52968  
ID ABP52968 standard; peptide; 22 AA.

XX  
AC ABP52968;  
XX  
DT 05-NOV-2002 (first entry)

XX  
DE Cellular proliferation inhibitor related peptide SEQ ID NO:3.

XX  
KW Cellular proliferation inhibition; cytostatic; anti-inflammatory; cancer;  
KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytal disorder; hypochalamic disorder; inflammatory;  
KW glandular disorder; macrophagal disorder; epithelial disorder;  
KW stromal disorder; blastocoelec disorder; angiogenic disorder;  
KW immunologic disorder.

XX  
OS Homo sapiens.

XX  
OS Mus musculus.

XX  
OS Synthetic.

XX  
PN WO200264779-A2.

XX  
PD 22-AUG-2002.

XX  
PF 21-JAN-2002; 2002WO-US001708.

XX  
PR 14-FEB-2001; 2001US-00782816.

XX  
PA (REGC ) UNIV CALIFORNIA.

XX  
PI Sharp DJ, Rogers GC, Scholey JM;  
XX  
DR WPI; 2002-657599/70.

XX  
PT New peptide inhibitors of p50/dynaminin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.

XX  
PS Claim 1; Page 29; 55pp; English.

XX  
The present invention describes an isolated peptide (I) comprising or  
CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
CC terminal extensions. (1) have cytostatic and anti-inflammatory activities  
CC and can be used as p50/dynaminin inhibitors and in gene therapy. The  
CC peptides, nucleic acid molecules and methods from the present invention  
CC are useful for treating cancer by inhibiting cellular proliferation, such  
CC as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and

CC neck tumours); leukaemias and lymphoid malignancies, other disorders such  
CC as neuronal, glial, astrocytal, hypothalamic and other glandular,  
CC macrophagal, epithelial, stromal and blastocoeleic disorders; and  
CC inflammatory, angiogenic and immunologic disorders. The present sequence  
CC represents a peptide that can be N-terminally added to (P1)

XX Sequence 22 AA;

Query Match 45.0%; Score 113; DB 5; Length 22;  
Best Local Similarity 100.0%; Pred. No. 3.4e-06;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVKETPOOKYORLLHEVOELTT 22  
Db 1 GVKETPOOKYORLLHEVOELTT 22

RESULT 10  
ABP52969  
ID ABP52969 standard; peptide; 21 AA.

XX AC ABP52969;

XX DT 05-NOV-2002 (first entry)

XX DE Cellular proliferation inhibitor related peptide SEQ ID NO:4.

XX KW Cellular proliferation inhibition; cytostatic; antiinflammatory; cancer;  
KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukaemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;  
KW glandular disorder; macrophagal disorder; epithelial disorder;  
KW stromal disorder; blastocoeleic disorder; angiogenic disorder;  
KW immunologic disorder.

XX OS Homo sapiens.  
OS Mus musculus.  
OS Synthetic.

XX PN WO200264779-A2.

XX PD 22-AUG-2002.

XX PF 21-JAN-2002; 2002WO-US001708.

XX PR 14-FEB-2001; 2001US-00782816.

XX PA (REGC ) UNIV CALIFORNIA.

XX PI Sharp DJ, Rogers GC, Scholey JM;

XX DR WPI; 2002-657599/70.

XX PT New peptide inhibitors of p50/dynaminin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.

XX PS Claim 1; Page 29; 55pp; English.

XX CC The present invention describes an isolated peptide (I) comprising or  
CC having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
CC sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
CC terminal extensions. (I) have cytostatic and antiinflammatory activities  
CC and can be used as p50/dynaminin inhibitors and in gene therapy. The  
CC peptides, nucleic acid molecules and methods from the present invention  
CC are useful for treating cancer by inhibiting cellular proliferation, such  
CC as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
CC gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
CC neck tumors); leukaemias and lymphoid malignancies, other disorders such  
CC as neuronal, glial, astrocytal, hypothalamic and other glandular,  
CC macrophagal, epithelial, stromal and blastocoeleic disorders; and

CC inflammatory, angiogenic and immunologic disorders. The present sequence  
CC represents a peptide that can be N-terminally added to (P1)

XX Sequence 21 AA;

Query Match 42.6%; Score 107; DB 5; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VKETPOOKYORLLHEVOELTT 22  
Db 1 VKETPOOKYORLLHEVOELTT 21

RESULT 11  
ABP52970  
ID ABP52970 standard; peptide; 20 AA.

XX AC ABP52970;

XX DT 05-NOV-2002 (first entry)

XX DE Cellular proliferation inhibitor related peptide SEQ ID NO:5.

XX KW Cellular proliferation inhibition; cytostatic; antiinflammatory; cancer;  
KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukaemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;  
KW glandular disorder; macrophagal disorder; epithelial disorder;  
KW stromal disorder; blastocoeleic disorder; angiogenic disorder;  
KW immunologic disorder.

XX OS Homo sapiens.  
OS Mus musculus.  
OS Synthetic.

XX PN WO200264779-A2.

XX PD 22-AUG-2002.

XX PF 21-JAN-2002; 2002WO-US001708.

XX PR 14-FEB-2001; 2001US-00782816.

XX PA (REGC ) UNIV CALIFORNIA.

XX PI Sharp DJ, Rogers GC, Scholey JM;

XX DR WPI; 2002-657599/70.

XX PT New peptide inhibitors of p50/dynaminin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.

XX PS Claim 1; Page 29; 55pp; English.

XX CC The present invention describes an isolated peptide (II) comprising or  
XX having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
XX sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
XX terminal extensions. (II) have cytostatic and antiinflammatory activities  
XX and can be used as p50/dynaminin inhibitors and in gene therapy. The  
XX peptides, nucleic acid molecules and methods from the present invention  
XX are useful for treating cancer by inhibiting cellular proliferation, such  
XX as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
XX gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
XX thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
XX neck tumors); leukaemias and lymphoid malignancies, other disorders such  
XX as neuronal, glial, astrocytal, hypothalamic and other glandular,  
XX macrophagal, epithelial, stromal and blastocoeleic disorders; and  
XX inflammatory, angiogenic and immunologic disorders. The present sequence  
XX represents a peptide that can be N-terminally added to (P1)



SQ Sequence 20 AA;  
Query Match 41.0%; Score 103; DB 5; Length 20;  
Best Local Similarity 100.0%; Pred. No. 4.5e-05;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 3 KETPOOKYORLHVEQELTT 22  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 KETPOOKYORLHVEQELTT 20

RESULT 12  
ABP52966  
ID ABP52966 standard; peptide, 23 AA.  
XX AC ABP52966;  
XX DT 05-NOV-2002 (first entry)  
XX DE Cellular proliferation inhibitor related peptide SEQ ID NO:1.  
XX KW Cellular proliferation inhibitor; cytostatic; antiinflammatory; cancer;  
KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;  
KW glandular disorder; macrophagal disorder; epithelial disorder;  
KW stromal disorder; blastocoelec disorder; angiogenic disorder;  
KW immunologic disorder.  
XX KW  
XX OS Homo sapiens.  
XX OS Mus musculus.  
XX OS Synthetic.  
XX FT Key Location/Qualifiers  
FT Misc-difference 22 /label= 'Leu, Val  
FT /note= "Leu in humans and Val in Mus musculus"  
XX PN WO200264779-A2.  
XX PD 22-AUG-2002.  
XX PF 21-JAN-2002; 2002WO-US001708.  
XX PR 14-FEB-2001; 2001US-00782816.  
XX PA (REGC ) UNIV CALIFORNIA.  
XX PI Sharp DJ, Rogers GC, Scholey JM;  
XX WPI; 2002-657599/70.  
XX PT New peptide inhibitors of p50/dynaminin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.  
XX PS Claim 1; Page 29; 55pp; English.

The present invention describes an isolated peptide (I) comprising or  
having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
terminal extensions. (1) have cytostatic and antiinflammatory activities  
and can be used as p50/dynaminin inhibitors and in gene therapy. The  
peptides, nucleic acid molecules and methods from the present invention  
are useful for treating cancer by inhibiting cellular proliferation, such  
as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
neck tumors); leukemias and lymphoid malignancies, other disorders such  
as neuronal, glial, astrocytal, hypothalamic and other glandular,  
macrophagal, epithelial, stromal and blastocoelec disorders; and  
inflammatory, angiogenic and immunologic disorders

XX SQ Sequence 23 AA;  
Query Match 41.0%; Score 103; DB 5; Length 23;  
Best Local Similarity 95.7%; Pred. No. 5.3e-05;  
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 23 EWEKIKTTVKESATBEKLPVLL 45  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 EWEKIKTTVKESATBEKLPVLL 23

RESULT 13  
ABP52971  
ID ABP52971 standard; peptide, 19 AA.  
XX AC ABP52971;  
XX DT 05-NOV-2002 (first entry)  
XX DE Cellular proliferation inhibitor related peptide SEQ ID NO:6.  
XX KW Cellular proliferation inhibitor; cytostatic; antiinflammatory; cancer;  
KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;  
KW glandular disorder; macrophagal disorder; epithelial disorder;  
KW stromal disorder; blastocoelec disorder; angiogenic disorder;  
KW immunologic disorder.  
XX KW  
XX OS Homo sapiens.  
XX OS Mus musculus.  
XX OS Synthetic.  
XX FT Key Location/Qualifiers  
FT Misc-difference 22 /label= 'Leu, Val  
FT /note= "Leu in humans and Val in Mus musculus"  
XX PN WO200264779-A2.  
XX PD 22-AUG-2002.  
XX PF 21-JAN-2002; 2002WO-US001708.  
XX PR 14-FEB-2001; 2001US-00782816.  
XX PA (REGC ) UNIV CALIFORNIA.  
XX PI Sharp DJ, Rogers GC, Scholey JM;  
XX WPI; 2002-657599/70.  
XX PT New peptide inhibitors of p50/dynaminin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.  
XX PS Claim 1; Page 29; 55pp; English.

The present invention describes an isolated peptide (I) comprising or  
having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
terminal extensions. (1) have cytostatic and antiinflammatory activities  
and can be used as p50/dynaminin inhibitors and in gene therapy. The  
peptides, nucleic acid molecules and methods from the present invention  
are useful for treating cancer by inhibiting cellular proliferation, such  
as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
neck tumors); leukemias and lymphoid malignancies, other disorders such  
as neuronal, glial, astrocytal, hypothalamic and other glandular,  
macrophagal, epithelial, stromal and blastocoelec disorders; and  
inflammatory, angiogenic and immunologic disorders. The present sequence  
represents a peptide that can be N-terminally added to (P1)

Query Match 39.0%; Score 96; DB 5; Length 19;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ETPQOKYORLLHVEVQLTT 22  
DB 1 ETPQOKYORLLHVEVQLTT 19

## RESULT 14

ABP52972 ID ABP52972 standard; peptide; 18 AA.

XX AC ABP52972;

DT 05-NOV-2002 (first entry)

XX Cellular proliferation inhibitor related peptide SEQ ID NO:7.

XX Cellular proliferation inhibition; cytostatic; antiinflammatory; cancer;  
KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukaemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;  
KW glandular disorder; macrophagal disorder; epithelial disorder;  
KW stromal disorder; blastocoealic disorder; angiogenic disorder;  
KW immunologic disorder.

XX OS Homo sapiens.  
OS Mus musculus.  
OS Synthetic.

XX WO200264779-A2.

XX 22-AUG-2002.

XX 21-JAN-2002; 2002WO-US001708.

XX 14-FEB-2001; 2001US-00782816.

XX (REGC ) UNIV CALIFORNIA.

XX Sharp DJ, Rogers GC, Scholey JM;

XX WPI; 2002-657599/70.

PT New peptide inhibitors of p50/dynaminin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.

XX Claim 1; Page 29; 55pp; English.

XX The present invention describes an isolated peptide (I) comprising or  
XX having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
XX sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
XX terminal extensions. (I) have cytostatic and antiinflammatory activities  
XX and can be used as p50/dynaminin inhibitors and in gene therapy. The  
XX peptides, nucleic acid molecules and methods from the present invention  
XX are useful for treating cancer by inhibiting cellular proliferation, such  
XX as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
XX gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
XX thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
XX neck tumors); leukemias and lymphoid malignancies, other disorders such  
XX as neuronal, glial, astrocytal, hypothalamic and other glandular,  
XX macrophagal, epithelial, stromal and blastocoealic disorders; and  
XX inflammatory, angiogenic and immunologic disorders. The present sequence  
XX represents a peptide that can be N-terminally added to (P1)

XX Sequence 18 AA;

Query Match 37.1%; Score 93; DB 5; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.00058;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TPQOKYORLLHVEVQLTT 22  
DB 1 TPQOKYORLLHVEVQLTT 18

## RESULT 15

ABP52973 ID ABP52973 standard; peptide; 17 AA.

XX AC ABP52973;

DT 05-NOV-2002 (first entry)

XX Cellular proliferation inhibitor related peptide SEQ ID NO:8.

XX Cellular proliferation inhibition; cytostatic; antiinflammatory; cancer;  
KW p50 inhibitor; dynaminin inhibitor; gene therapy; tumour; carcinoma;  
KW sarcoma; glioblastoma; leukaemia; lymphoid malignancy; neuronal disorder;  
KW glial disorder; astrocytal disorder; hypothalamic disorder; inflammatory;  
KW glandular disorder; macrophagal disorder; epithelial disorder;  
KW stromal disorder; blastocoealic disorder; angiogenic disorder;  
KW immunologic disorder.

XX OS Homo sapiens.  
OS Mus musculus.  
OS Synthetic.

XX WO200264779-A2.

XX 22-AUG-2002.

XX 21-JAN-2002; 2002WO-US001708.

XX 14-FEB-2001; 2001US-00782816.

XX (REGC ) UNIV CALIFORNIA.

XX Sharp DJ, Rogers GC, Scholey JM;

XX WPI; 2002-657599/70.

PT New peptide inhibitors of p50/dynaminin useful for treating cancer by  
PT inhibiting cellular proliferation, e.g. benign or malignant tumors,  
PT leukemia and lymphoid malignancies, or inflammatory, angiogenic and  
PT immunologic disorders.

XX Claim 1; Page 29; 55pp; English.

XX The present invention describes an isolated peptide (I) comprising or  
XX having at least 90% identity to (P1) or (P2). Where (P1) and (P2) are the  
XX sequences given in ABP52966 and ABP52967 and can have C-terminal and N-  
XX terminal extensions. (I) have cytostatic and antiinflammatory activities  
XX and can be used as p50/dynaminin inhibitors and in gene therapy. The  
XX peptides, nucleic acid molecules and methods from the present invention  
XX are useful for treating cancer by inhibiting cellular proliferation, such  
XX as benign or malignant tumors (renal, liver, kidney, bladder, breast,  
XX gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,  
XX thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and  
XX neck tumors); leukemias and lymphoid malignancies, other disorders such  
XX as neuronal, glial, astrocytal, hypothalamic and other glandular,  
XX macrophagal, epithelial, stromal and blastocoealic disorders; and  
XX inflammatory, angiogenic and immunologic disorders. The present sequence  
XX represents a peptide that can be N-terminally added to (P1)

XX Sequence 17 AA;

Query Match 35.1%; Score 88; DB 5; Length 17;  
Best Local Similarity 100.0%; Pred. No. 0.0021;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PQQKYORLLHVEVQLTT 22  
|||||



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2005, 21:43:09 / Search time 18.9091 Seconds  
(without alignments)  
264.596 Million cell updates/sec

Title: 09782816-3-1-22

Sequence: 1 GVKETPOQKQRLHVEVQEL.....ESATREKLTPLVLAKQALAL 52

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79.\*  
1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69.5	27.7	804	2 T132864	hypothetical prote
2	68	27.1	1742	2 T49451	kinesin-like prote
3	67.5	26.9	1325	2 S16129	dynein-associated
4	67	26.7	1287	2 T22235	hypothetical prote
5	67	26.7	2481	2 D90011	hypothetical prote
6	66	26.3	393	2 T04957	FM2B protein (lipo
7	66	26.3	866	2 C71509	hypothetical prote
8	65.5	26.1	1837	2 T41023	probable DNA polym
9	64	25.5	210	2 B84499	probable nuclear p
10	63.5	25.3	163	2 S69518	hypothetical prote
11	63.5	25.3	374	2 D70391	hypothetical prote
12	63	25.1	281	2 F75216	hypothetical prote
13	63	25.1	314	2 F90577	lipoprotein vsa
14	62.5	24.9	1281	2 JCS368	dynactin 1 - mouse
15	62	24.7	296	2 T22232	hypothetical prote
16	62	24.7	2469	2 H36812	hypothetical prote
17	61	24.3	442	2 C95070	hypothetical prote
18	61	24.3	442	2 A97938	sensor histidine k
19	60.5	24.1	819	2 AC2029	vmsd, histidine ki
20	60.5	24.1	1214	2 JC2069	hypothetical prote
21	60	23.9	523	2 S06920	zinc-finger protei
22	60	23.9	695	2 H72243	glucose transport
23	59.5	23.7	279	2 D71453	translation elonga
24	59.5	23.7	665	2 S62328	hypothetical prote
25	59	23.5	112	2 G72645	kinesin-like DNA b
26	59	23.5	239	2 D84004	hypothetical prote
27	59	23.5	550	2 G85436	hypothetical prote
28	59	23.5	611	2 B90543	glucose inhibited
29	59	23.5	782	2 A34219	Bic-D protein - fr

30	59	23.5	951	2 T45726	hypothetical prote
31	59	23.5	2101	2 A42184	nuclear mitotic ap
32	58.5	23.3	154	2 T25602	hypothetical prote
33	58.5	23.3	821	2 A12417	hypothetical prote
34	58	23.1	302	2 T01615	hypothetical prote
35	58	23.1	319	2 T25666	hypothetical prote
36	58	23.1	623	2 A48315	hypothetical prote
37	58	23.1	630	2 D97992	lamin III - Africa
38	58	23.1	677	2 T22333	hypothetical prote
39	58	23.1	707	2 S78538	histidine ammonia-
40	58	23.1	727	2 AD2188	site-specific reco
41	58	23.1	740	2 G95153	hypothetical prote
42	58	23.1	978	2 A70387	neuraminidase, pro
43	58	23.1	2139	2 T18286	conserved hypotet
44	58	23.1	2422	2 T12687	myosin heavy chain
45	57.5	22.9	237	2 C87656	AKR protein homolo
					GDDEF family prote

## ALIGNMENTS

RESULT 1  
T132864  
hypothetical protein ZK484.4 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T132864  
R/Murray, J.; Wohldmann, P.  
submitted to the EMBL Data Library, December 1997  
A/Description: The sequence of C. elegans cosmid ZK484.  
A/Reference number: Z21237  
A/Accession: T132864  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Reidues: 1-804 <MUR>  
A/Cross-references: UNIPROT:Q95X56; EMBL:AF040659; PDB:AA095062.1; GSPDB:GN00019; CESP:  
A/Experimental source: strain Bristol N2; clone ZK484  
C/Genetics:  
A/Gene: CESP:ZK484.4  
A/Map position: 1  
A/Intons: 52/1; 113/1; 162/3; 201/3; 248/2; 346/1; 433/2; 464/3; 565/3; 628/2; 646/3; 7  
C/Superfamily: Caenorhabditis elegans hypothetical protein ZK484.4

Query Match 27.7%; Score 69.5; DB 2; Length 804;  
Best Local Similarity 34.4%; Pred. No. 20;  
Matches 22; Conservative 10; Mismatches 19; Indels 13; Gaps 3;

Oy 2 VKETPOQKQRLHVEVQELTVE-----KIKT-TVESATRE-KLTPVLAKQ 48  
Db 366 VEDPDERKYQKTKTVKEIENNEDEGEIGASTSKRKTIVIKKEIDEVKOTPAKURA 425

Oy 49 LAAL 52  
Db 426 SAKL 429

RESULT 2  
T49451  
kinesin-like protein Kif21a related protein [imported] - Neurospora crassa  
N/Alternate names: protein B14D6.30  
C/Species: Neurospora crassa  
C/Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 09-Jul-2004  
R/Schulte, U.; Algen, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, May 2000  
A/Reference number: Z25022  
A/Accession: T49451  
A/Status: preliminary  
A/Molecule type: DNA  
A/Reidues: 1-1742 <SCH>  
A/Cross-references: UNIPROT:Q8X0C5; EMBL:AL356173; GSPDB:GN00116; NCSP:B14D6.30  
A/Experimental source: BAC clone B14D6; strain OR74A  
C/Genetics:

A:Gene: NCSP:BIAD6.30  
A:Map position: 6  
A:Introns: 58/1; 166/1; 267/3; 1543/3

Query Match	27.1%	Score 68	DB 2	Length 1742
Best Local Similarity	36.6%	Pred. No. 66		
Matches 15, Conservative		8, Mismatches	18, Indels	0, Gaps 0;

Oy 7 QOKYORLHEVOELTTEVEKIKITTKESATEEKLTPVLLAK 47  
::|::|::|::|::|::|::|  
Db 950 EEKHQETLDMVELKTEIAKAQAALSVESSISRTSTPVIRRK 990

RESULT 3  
S16129

dynein-associated protein, 150K, cytosolic - bovine  
 C|Species: Bos primigenius taurus (cattle)  
 C|Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
 C|Accession: S16129  
 R|Holzhauer, E.L.F.; Hammarback, J.A.; Paschal, B.M.; Kravitt, N.G.; Pfister, K.K.; Valleer  
 Nature 351, 579-583, 1991  
 A|Title: Homology of a 150K cytoplasmic dynein-associated polypeptide with the Drosophila  
 A|Reference number: S16129; MUID:91260877; PMID:1828535  
 A|Accession: S16129  
 A|Status: preliminary  
 A|Molecule type: mRNA  
 A|Residues: 1-1325 <KTV>  
 A|Cross-references: UNIPROT:P28023

Query Match	26.9%	Score	67.5	DB	2	Length	1325
Best Local Similarity	34.7%	Pred.	No. 56				
Matches	17	Conservative	11	Mismatches	20	Indels	1
						Gaps	1

```

Oy      1 GVKETPOOKRÖRLNHEVOELTTEVEKIKTTVKESATEBKLTPLVILAKÖL 49
        | : | : | | | : : : : : : : | | |
Db      314 GAESLÖGEVALKERVDELTTDLLEILKAIEBKSGDGAASÖYL-KÖL 361

```

## RESULT 4

hypochemical protein F45G2.3 - *Caenorhabditis elegans*  
 C|Species: *Caenorhabditis elegans*  
 C|Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C|Accession: T22235  
 R|Lindsay, S.  
 submitted to the EMBL Data Library, March 1997  
 A|Reference number: Z19535  
 A|Accession: T22235  
 A|Status: preliminary; translated from GB/EMBL/DBJ  
 A|Molecule type: DNA  
 A|Residues: 1-1287 (#MW)  
 A|Cross-references: UNIPROT:062245; EMBL:593382; PIDD:CAM07612.1; GSPDB:GN00021; CESP:F45G2  
 A|Experimental source: clone F45G2  
 C|Gene: CESP:F45G2.3  
 A|Map position: 3  
 A|Introms: 59/3; 117/3; 153/3; 180/2; 233/2; 341/3; 393/1; 424/2; 551/3; 664/3; 734/3; 8

```

Query Match 26.7%; Score 67; DB 2; Length 1287;
Best Local Similarity 31.6%; Pred. No. 61;
Matches 18; Conservative 11; Mismatches 16; Indels 10; Gaps 2;

Cy 4 ETPQKRYRLHVEQVLTTEV--EKIKTTVESATEBK-----TPVLIANQLA 50
Db 1122 EPPPKRKRIILEVIDDSEFFQEKYKENTKESVAEKEKENSNEELKDPPIIRKAVS 1178

```

## RESULT 5

FmtB protein (imported) - Staphylococcus aureus (strain N315)  
 C:Species: Staphylococcus aureus  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #ext\_change 09-Jul-2004  
 C:Accession: D90011

R. Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguci, ma, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; I. C.; Shida, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001

ArTitle: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.

A/Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*  
A/Reference number: A89758; NCID:21311952; PMID:11418146  
A/Accession: D90011

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 12481 <RD>  
A:Cross-references: UNIPROT:Q99QH6; GB:BA000018; P1D:G13701961; P1DN:BAB43253.1; GSPDB:G  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: fmcB (mrp)

Query Match	26.7%;	Score 67;	DB 2;	Length 2481;
Best Local Similarity	35.1%;	Pred. No. 1.2e+02;		
Matches 13; Conservative		9; Mismatches 15; Indels		

```
Oy      2 VKETPOQKOPRLHVOELTTEVEKIKTYKESATBE 38
          ::||| : |||::|||::|||
Db      1609 IEQTNNAQQEINDAKQEAVTELINAKTNVDQSSTNE 1645
```

RESULT 6  
T04957

hypothetical protein F7J7\_200 - *Arabidopsis thaliana*  
C/Species: *Arabidopsis thaliana* (mouse-ear cress)  
C/Date: 23-Apr-1999 #sequence\_rev1999 #text\_change 09-Jul-2004  
C/Accession: T04957  
R/Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X  
submitted to the Protein Sequence Database, July 1998  
A/Reference number: Z15391  
A/Accession: T04957  
A/Molecule type: DNA  
A/Residues: 1-393 <BEV>  
A/Cross-References: UNIPROT:O49567; EMBL:AL021960  
A/Experimental source: cultivar Columbia; BAC clone F7J7  
C/Genetics:  
A/Map position: 4  
A/Introns: 80/1, 217/3, 264/1, 291/3  
A/Note: F7J7\_200

Query Match	26.3%	Score 66	DB 2	Length 393
Best Local Similarity	38.6%	Pred. No. 21		
Matches 17	Conservative 12	Mismatches 7	Indels 8	Gaps 2

```
Oy      3 KETPOOKORLLHEVQELTTEVEKIKTTVESATEE-KLTPVL 45
          ::||::|::| |::|::| |::|::| :
Db     27 QESTQGTHQNLLYKVQ-----KWRTSLKDSSDAELKLSPALV 63
```

## RESULT 7

probable DNA polymerase I - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
 C:Species: Chlamydia trachomatis  
 C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 09-Jul-2004  
 C:Accession: C71509  
 R:Stehenson, R.S.; Kalmun, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998  
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis  
 A:Reference number: A71570; NCID:9900809; PMID:9784116  
 A:Accession: C71509  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-866 <ARN>  
 A:Cross-references: UNIPROT:O84500; GB:AE00132; GB:AE001273; NID:g3328916; PIDN:AAC6809  
 A:Experimental source: serotype D, strain UW-3/Cx  
 C:Genetics:  
 A:Gene: polA  
 C:Superfamily: DNA-directed DNA polymerase I

Best Local Similarity 36.4%; Pred. No. 50;  
Matches 20; Conservative 10; Mismatches 21; Indels 4; Gaps 2;

Qy 1 GVKETPOOKYORLLHE---VOELTTEVEKIKITTVESATEEKLTPVLLAKQIAAL 52  
Db 188 GVSQSGPFGKQAALKEFQSVELVANTRERLSGKTKQWIEDQKET-IILSKRLATL 241

## RESULT 8

T41023  
probable nuclear pore complex-associated protein - fission yeast (Schizosaccharomyces po

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C:Accession: T41023

R:Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, June 1998

A:Reference number: Z21965

A:Accession: T41023

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1837 <MCR>

A:Cross-references: UNIPROT:O74424; EMBL:AL023860; PIDN:CAA19588.1; GSPDB:GN00068; SPDB:

A:Experimental source: strain 972h-; cosmid c162

C:Genetics:

A:Gene: SPDB:SPC162.08C

A:Map position: 3

Query Match 26.1%; Score 65.5; DB 2; Length 1837;  
Best Local Similarity 35.8%; Pred. No. 1.3e+02;

Matches 19; Conservative 10; Mismatches 17; Indels 7; Gaps 2;

Qy 7 QOKYORLLHEVOELTTEVEKIKITTV--KESATEEKLTPVLLAK-----QIAL 52  
Db 662 QTSYORSLIRLEQLTNLESLKISRNKEKKEEALISLOLEKSNIOQLTSL 714

## RESULT 9

E84499  
hypothetical protein At2g11890 (imported) - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 12-Jul-2004

C:Accession: E84499

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffit, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.

euser, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617137

A:Accession: E84499

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-210 <STO>

A:Cross-references: UNIPROT:Q9S1Y3; GB:AB002093; NID:94557062; PIDN:AAD2501.1; GSPDB:GN

C:Genetics:

A:Gene: At2g11890

A:Map position: 2

C:Superfamily: uncharacterized CYTH domain protein

Query Match 25.5%; Score 64; DB 2; Length 210;  
Best Local Similarity 40.5%; Pred. No. 17;

Matches 17; Conservative 9; Mismatches 14; Indels 2; Gaps 2;

Qy 1 GVK-ETPOOKYQ-RLLHEVOELTTEVEKIKITTVESATEEKL 40  
Db 150 GVKLEVDETKDFGNCYIEICETSEPERVKTMIEFLTEK 191

## RESULT 10

S69518  
hypothetical protein 12 - phage HPI

C:Species: phage HPI

C:Date: 06-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004

C:Accession: S69518

R:Esposito, D.; Fitzmaurice, W.P.; Benjamin, R.C.; Goodman, S.D.; Waldman, A.S.; Scocca,  
Nucleic Acids Res. 24, 2360-2368, 1996

A:Title: The complete nucleotide sequence of bacteriophage HPI DNA.

A:Reference number: S69503; MUID:96279738; PMID:8710508

A:Accession: S69518

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-163 <ESP>

A:Cross-references: UNIPROT:P51714; EMBL:U24159; NID:g1046235; PIDN:AAB09197.1; PID:g104

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995

C:Superfamily: phage HPI hypothetical protein 12

Query Match 25.3%; Score 63.5; DB 2; Length 163;  
Best Local Similarity 37.8%; Pred. No. 15;

Matches 17; Conservative 11; Mismatches 16; Indels 1; Gaps 1;

Qy 5 TPQKQYORLLHEVOELTTEVEKIK-TTVKESATEEKLTPVLLAKQ 48  
Db 25 TPQKQYKLMHEFELCSGVAKKRPVYIKSDIGDPCFVVMVILAKQ 69

## RESULT 11

D70391  
hypothetical protein aq\_1060 - Aquifex aeolicus

C:Species: Aquifex aeolicus

C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004

C:Accession: D70391

R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov-

V. Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70300; MUID:98196666; PMID:9537320

A:Accession: D70391

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-374 <AOF>

A:Cross-references: UNIPROT:O67159; GB:AB000721; NID:g2983544; PIDN:AAC07129.1; PID:g298

A:Experimental source: strain VFS

C:Genetics:

A:Gene: aq\_1060

C:Superfamily: multidrug resistance protein A; 11poy1/biotin-binding homology

Query Match 25.3%; Score 63.5; DB 2; Length 374;  
Best Local Similarity 31.1%; Pred. No. 36;

Matches 19; Conservative 14; Mismatches 13; Indels 15; Gaps 3;

Qy 6 PQQYKQK-----LHBE-----VOELTTEVEKIKITTVESATEEKLTPVLLAKQIAA 51  
Db 164 PRKFEVDVNLKVLHREYLEKSIQIEINTEIKRAKGI-ENARNKFKTIEELKEKLS 222

## RESULT 12

F75216

hypothetical protein PAB2181 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi

C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004

C:Accession: F75216

R:anonymous, Genomecope

submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru-

A:Reference number: A75001

A:Accession: F75216

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-281 <KAW>

A:Cross-references: UNIPROT:Q9V217; GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB4918

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB2181

Query Match 25.1%; Score 63; DB 2; Length 281;  
Best Local Similarity 36.6%; Pred. No. 30;  
Matches 15; Conservative 12; Mismatches 12; Indels 2; Gaps 2;

Qy 1 GVKE-TPQOKYQRLHVEQELTTEVEKIKTVKESATEEKL 40  
Db 119 GKEVVAEEVEKELKKEYEKLKQFEFEVKAKI-EAAELES 158

RESULT 13  
F90577  
lipoprotein vaab [imported] - Mycoplasma pulmonis (strain UAB CTIP) (fragment)

C/Species: Mycoplasma pulmonis  
C/Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
C/Accession: F90577  
R/Chambaud, I.; Heilbr, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;  
Nucleic Acids Res. 29, 2145-2153, 2001  
A/Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pult  
A/Reference number: A99512; MUID:21267165; PMID:11353084  
A/Accession: F90577

A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-314 <KUN>  
A/Cross-references: UNIPROT:Q98Q43; GB:AL445566; PID:g14089940; PIDN:CAC13699.1; GSPDB:C  
A/Experimental source: strain UAB CTIP  
C/Genetics:

A/Gene: MYPu\_5260  
A/Genetic code: SGC3

Query Match 25.1%; Score 63; DB 2; Length 314;  
Best Local Similarity 29.8%; Pred. No. 34;  
Matches 14; Conservative 17; Mismatches 10; Indels 6; Gaps 2;

Qy 2 VKETPQOKYQRLHVEQELTTEVEKIKTVKESATEEKL 44  
Db 246 ITTEKAKFKVDVIOPARTKQDLTKLEKIKS-EKENIEKKLDP 290

## RESULT 14

JC5368  
dynamitin 1 - mouse  
N/Alternate names: p150 Glued  
C/Species: Mus musculus (house mouse)  
C/Date: 28-May-1997 #sequence\_revision 18-Jul-1997 #text\_change 09-Jul-2004  
C/Accession: JC5368

R/Jiang, W.; Weber, J.S.; Tokito, M.K.; Holzbaun, E.L.F.; Weisler, M.H.  
Biochem. Biophys. Res. Commun. 231, 344-347, 1997

A/Title: Mouse p150Glued (dynamitin 1) cDNA sequence and evaluation as a candidate for th  
A/Reference number: JC5368; MUID:97223454; PMID:9070275

A/Accession: JC5368  
A/Molecule type: mRNA

A/Residues: 1-1281 <UN>

A/Cross-references: UNIPROT:O08788; GB:U60312; NID:g2104494; PIDN:AAB5773.1; PID:g21044

A/Experimental source: brain  
C/Comment: This protein is a member of the oligomeric dynamitin complex that is required

C/Genetics:

A/Gene: Dctn1

A/Map position: 6

Query Match 24.9%; Score 62.5; DB 2; Length 1281;  
Best Local Similarity 34.8%; Pred. No. 1.8e+02;  
Matches 16; Conservative 11; Mismatches 18; Indels 1; Gaps 1;

Qy 4 ETPQOKYQRLHVEQELTTEVEKIKTVKESATEEKLTPVLI 49  
Db 319 ESLQOEVALKRVDELTTDLIELKAIIEKSGDGAASSYQL-KQL 363

## RESULT 15

T32222  
hypochemical protein T23B12.7 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T32222

R/Davidson, S.; Wohldmann, P.; Gillam, B.

submitted to the EMBL Data Library, September 1997

A/Description: The sequence of C. elegans cosmid T23B12.

A/Reference number: Z21137

A/Accession: T32222

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-296 <DAV>

A/Cross-references: UNIPROT:O17002; EMBL:AF022982; PIDN:AAB69938.1; GSPDB:GN00023; CESP:

A/Experimental source: strain Bristol N2; clone T23B12

C/Genetics:

A/Gene: CESP:T23B12.7

A/Map position: 5

A/Introns: 67/3

Query Match 24.7%; Score 62; DB 2; Length 296;  
Best Local Similarity 31.0%; Pred. No. 40;  
Matches 18; Conservative 12; Mismatches 12; Indels 16; Gaps 3;

Qy 3 KETPQOKYQRLHVEQELTTEVEK-ITTVKESAT-----EKLTPVLI 47  
Db 132 KKTDOIRNFK--EIEAIRQOLEKEVNEVYKQATLMKTEREKHQKQOEKLTTPRL 186

Search completed: November 3, 2005, 22:04:16  
Job time : 20.9091 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2005, 21:37:29 ; Search time 91.4866 Seconds

(without alignments)  
291.060 Million cell updates/sec

Title: 09782816-3-1-22

Perfect score: 251

Sequence: 1 GVKEETPOOKTQRLHVEQEL.....ESATEBKLTPLVLAQKLAAL 52

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.\*

1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	251	100.0	400	1	DCT2_HUMAN
2*	248	98.8	401	1	DCT2_MOUSE
3	248	98.8	402	1	O6A3H5
4	220	87.6	403	2	O6IRB3
5	213	84.9	403	2	O66J30
6	211	84.1	338	2	O7ZXV2
7	177	70.5	402	2	O9PTG6
8	177	70.5	405	2	O7T3H1
9	76	30.3	402	2	O6IP53
10	72.5	28.9	380	2	O9V4Y9
11	71	28.3	800	2	O6MVP7
12	71	28.3	813	2	O7SH14
13	70	27.9	311	2	O9TIV5
14	70	27.9	751	2	O8MOK1
15	70	27.9	1087	2	O8BBS5
16	70	27.9	1124	2	O8MOK8
17	70	27.9	2199	2	O7PUP2
18	69.5	27.7	873	2	O9SXS6
19	69.5	27.7	919	2	O9SXS5
20	68	27.1	1755	2	O7RVJ1
21	68	27.1	1968	2	O8XOC5
22	67.5	26.9	1868	2	O6BNV2
23	67	26.7	455	2	O7ZVF1
24	67	26.7	639	2	O6Z245
25	67	26.7	1795	2	O9LCI3
26	67	26.7	2478	2	O9LCI2
27	67	26.7	2481	2	O9RL69
28	67	26.7	2481	2	O99OR6
29	66	26.3	393	2	O7A4B1
30	66	26.3	393	2	O49567
31	66	26.3	860	2	O70009

32	66	26.3	866	2	O84500
33	66	26.3	1022	2	O8TBY8
34	65.5	26.1	1837	2	O74424
35	65	25.9	388	2	O7P225
36	64.5	25.7	348	2	O8LI79
37	64	25.5	149	2	O8B253
38	64	25.5	188	2	O9UM00
39	64	25.5	188	2	O92113
40	64	25.5	188	2	O6DGM9
41	64	25.5	210	2	O9SIY3
42	64	25.5	230	2	O75545
43	64	25.5	1227	1	J1P3.DROME
44	64	25.5	1956	2	O9Y2K3
45	63.5	25.3	163	1	Y012_BPHPI

## ALIGNMENTS

RESULT 1  
DCT2\_HUMAN  
ID DCT2\_HUMAN STANDARD: PRT, 400 AA.  
AC O13561; O86YN2; O9BW17;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 05-JUL-2004 (Rel. 44, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Dynactin complex 50 kDa subunit (50 kDa dynein-associated polypeptide)  
DE (p50 dynactin) (DCTN-50) (Dynactin 2).  
GN Name=DCTN2; Synonyms=DCTN50;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OC NCBI\_TaxID=9606;  
RX [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=96178072; PubMed=8647893; DOI=10.1083/jcb.132.4.617;  
RA Echeverri C.J., Paschal B.M., Vaughan K.T., Vallee R.B.;  
RT "Molecular characterization of the 50-kD subunit of dynactin reveals  
RT function for the complex in chromosome alignment and spindle  
RT organization during mitosis."  
RL J. Cell Biol. 132:617-633(1996).  
[2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta, Skin, and Uterus;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Strausberg R.L., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rabin G.M., Hong L.,  
RA Stappleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ueding T.B., Toshitsuki S., Carninci P., Prange C.,  
RA Rana S.S., Loguelfano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny K.S., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[3]  
RP SEQUENCE OF 1-384 FROM N.A.  
RA Annals J.P., Yu-Lee L.-Y.;  
RT "Human 50 kD dynactin subunit, p50 dynactin, isolated from HeLa  
RL cells."  
RN Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
[4]  
RP SEQUENCE OF 1-13.

CC TISSUE=Platelet;  
RX MEDLINE=22608298; PubMed=12665801; DOI=10.1038/nbr810;  
RA Gevaert K., Goethals M., Martens L., Van Damme J., Stees A.,  
RA Thomas G.R., Vandekerckhove J.,  
RT "Exploring proteomes and analyzing protein processing by mass  
RT spectrometric identification of sorted N-terminal peptides";  
RL Nat. Biotechnol. 21:566-569(2003).  
CC -1- FUNCTION: Modulates cytoplasmic dynein binding to an organelle,  
CC and plays a role in prometaphase chromosome alignment and spindle  
CC organization during mitosis. May play a role in synapse formation  
CC during brain development.  
CC -1- SUBUNIT: Subunit of dynein, a multiprotein complex associated  
CC with dynin.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL, U50733; AAC0423.1; -;  
DR EMBL, BC000718; AAH00718.1; -;  
DR EMBL, BC009468; AAH09468.1; -;  
DR EMBL, BC014083; AAH14083.1; -;  
DR EMBL, AY189155; AAC034395.1; -;  
DR Genbank, HGNC:2712; DCTN2.  
DR MIM, 607376; -;  
DR GO, GO:0005813; C:centrosome; TAS.  
DR GO, GO:0005869; C:dynein complex; TAS.  
DR GO, GO:0000776; C:kinetochore; TAS.  
DR GO, GO:0008283; P:cell proliferation; TAS.  
DR GO, GO:0007067; P:mitosis; TAS.  
DR InterPro, IPR006996; Dyneinlin.  
DR Pfam, PF04912; Dyneinlin; 1.  
KW Coiled coil; Cytoskeleton; Direct protein sequencing; Dynein;  
KW Membrane; Microtubule; Motor protein.  
FT INIT MET 0  
FT DOMAIN 98 131 Coiled coil (Potential).  
FT DOMAIN 213 243 Coiled coil (Potential).  
FT DOMAIN 378 398 Coiled coil (Potential).  
FT CONFLICT 34 34 A -> APAQL (in Ref. 1).  
FT CONFLICT 35 35 E -> ELR (in Ref. 3).  
FT CONFLICT 381 384 LATV -> PGHS (in Ref. 3).  
SQ SEQUENCE 400 AA; 44099 MW; 0A95AB95C0B8270F CRC64;  
  
Query Match 100.0%; Score 251; DB 1; Length 400;  
Best Local Similarity 100.0%; Pred. No. 4.7e-17;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 GVKETPOOKRQRLHVEVLTTEVEKITTVEKSTTEKTLPPVLAQOLAL 52  
Db 93 GVKETPOOKRQRLHVEVLTTEVEKITTVEKSTTEKTLPPVLAQOLAL 144  
  
RESULT 2  
DCT2\_MOUSE  
ID DCT2\_MOUSE STANDARD; PRT; 401 AA.  
AC Q99KJ8;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 05-JUN-2004 (Rel. 44, Last sequence update)  
DT 05-JUN-2004 (Rel. 44, Last annotation update)  
DE Dynein complex 50 kDa subunit (50 kDa dynein-associated polypeptide)  
DE (p50 dynein) (DCTN-50) (Dyactin 2) (Growth cone membrane protein  
DE 23-48K) (GMP23-48K).  
GN Name=Dctn2;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,  
RA Dictchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stadelton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Garinici P., Prange C.,  
RA Raha S.S., Loughlano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.U., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butcherfield Y.S.N., Krzywinski M.I., Skalek A., Smallus D.E.,  
RA Scherch A., Schein J.B., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE OF 65-74; 77-91; 102-116; 156-170; 194-216 AND 309-320,  
RP SUBCELLULAR LOCATION, AND DEVELOPMENTAL STAGE.  
RP TISSUE=Brain;  
RC MEDLINE=97289622; PubMed=9144527; DOI=10.1006/dbnc.1997.6447;  
RX Abe T.K., Tanaka H., Iwanga T., Odani S., Kuwano R.;  
RT "The presence of the 50-kDa subunit of dynein complex in the nerve  
RT growth cone";  
RL Biochem. Biophys. Res. Commun. 233:295-299(1997).  
RN [3]  
RP INTERACTION WITH BICD2.  
RX MEDLINE=21376052; PubMed=11483508; DOI=10.1093/emboj/20.15.4041;  
RA Hoogenraad C.C., Akhmanova A., Howell S.A., Dordland B.R.,  
RA de Zeeuw C.I., Willemsen R., Visser P., Grosved F., Galjart N.;  
RT "Mammalian Golgi-associated Bicaudal-D2 functions in the dynein-  
RT dynein pathway by interacting with these complexes";  
RL EMBO J. 20:4041-4054(2001).  
CC -1- FUNCTION: Modulates cytoplasmic dynein binding to an organelle,  
CC and plays a role in prometaphase chromosome alignment and spindle  
CC organization during mitosis. May play a role in synapse formation  
CC during brain development.  
CC -1- SUBUNIT: Subunit of dynein, a multiprotein complex associated  
CC with dynactin (By similarity). Interacts with BICD2.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.  
CC -1- DEVELOPMENTAL STAGE: Present at high levels in both cytoplasmic  
CC and membrane-associated forms in neonates. Levels of membrane-  
CC associated form are greatly reduced in the adult.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL, BC004613; AAH04613.1; -;  
DR MGD, MGI:107733; Dctn2.  
DR InterPro, IPR006996; Dyneinlin.  
DR Pfam, PF04912; Dyneinlin; 1.  
KW Coiled coil; Cytoskeleton; Direct protein sequencing; Dynein;  
KW Membrane; Microtubule; Motor protein.  
FT INIT MET 0  
FT DOMAIN 98 131 Coiled coil (Potential).  
FT DOMAIN 214 244 Coiled coil (Potential).  
SQ SEQUENCE 401 AA; 43985 MW; 1535E4ABD5940EBC CRC64;  
  
Query Match 98.8%; Score 248; DB 1; Length 401;  
Best Local Similarity 98.1%; Pred. No. 9.4e-17;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GVKETPOOKORLHBEVOELTTEVEKIKTYKESATBEKLTPLVLAQOLAL 52  
 DB 93 GVKETPOOKORLHBEVOELTTEVEKIKTYKESATBEKLTPLVLAQOLAL 144  
 RESULT 3  
 ID 06AYH5 PRELIMINARY; PRT; 402 AA.  
 AC 06AYH5; 25-OCT-2004 (TREMBlrel. 28, Created)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
 DE 25-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DE Dynactin 2.  
 GN Name=Dctn2;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative.";  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX Klein S., Strausberg R.;  
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC070987; AAH70987.1;  
 DR GO: GO:0005869; C:dynactin complex; IEA.  
 DR GO: GO:0007017; P:microtubule-based process; IEA.  
 DR InterPro: IPR006996; P:microtubule-based process; IEA.  
 DR Pfam: PF04912; Dynactin; 1.  
 SQ SEQUENCE 403 AA; 44148 MW; 55033553A4FB052 CRC64;  
 OY Query Match 98.8%; Score 248; DB 2; Length 402;  
 Best Local Similarity 98.1%; Pred. No. 9.4e-17;  
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 DB 94 GVKETPOOKORLHBEVOELTTEVEKIKTYKESATBEKLTPLVLAQOLAL 145  
 RESULT 4  
 ID 06IRB3 PRELIMINARY; PRT; 403 AA.  
 AC 06IRB3; 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DE 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Dctn2-prov protein.  
 GN Name=Dctn2-prov;  
 OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative.";  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX Klein S., Strausberg R.;  
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC070987; AAH70987.1;  
 DR GO: GO:0005869; C:dynactin complex; IEA.  
 DR GO: GO:0007017; P:microtubule-based process; IEA.  
 DR InterPro: IPR006996; P:microtubule-based process; IEA.  
 DR Pfam: PF04912; Dynactin; 1.  
 SQ SEQUENCE 403 AA; 44828 MW; 9FD158C7C7983062 CRC64;  
 OY Query Match 87.6%; Score 220; DB 2; Length 403;  
 Best Local Similarity 84.6%; Pred. No. 6e-14;  
 Matches 44; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
 DB 94 GVKETPOOKORLHBEVOELTTEVEKIKTYKESATBEKLTPLVLAQOLAL 145  
 RESULT 5  
 ID 06GJ30 PRELIMINARY; PRT; 403 AA.  
 AC 06GJ30; 25-OCT-2004 (TREMBlrel. 28, Created)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
 DE 25-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DE MGC82128 protein.  
 GN Name=MGC82128;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]

RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.;  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT initiative.";  
RL Dev. Dyn. 225:384-391(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shemmer C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RA Klein S., Gerhard D.S.;  
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC081081; AAB81081.1; -  
DR InterPro: IPR006996; Dynamitin.  
DR Pfam: PF04912; Dynamitin; 1.  
SQ SEQUENCE 403 AA; 44737 MW; 86BE8CEB54325EF3 CRC64;  
Query Match 84.9%; Score 213; DB 2; Length 403;  
Best Local Similarity 82.7%; Pred. No. 3e-13;  
Matches 43; Conservative 5; Mismatches 4; Indels 0; Gaps 0;  
OY 1 GVKETPOQKRYORLHVEQELTTEVEKIKTTVKESATFEKLTPLVALAKQVAL 52  
DB 94 GKMETPOQKRYORLHVEQELTTEVEKIKTTVKESATFEKLTPLVALAKQVAL 145  
RESULT 6  
O7ZXY2 PRELIMINARY; PRT; 338 AA.  
AC O7ZXY2;  
DT 01-JUN-2003 (TRENBLrel. 24, Created)  
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Dcch2-prov protein.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
NCBI\_TaxID=8355;  
OX [1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Embryo;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shemmer C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.;  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT initiative.";  
RL Dev. Dyn. 225:384-391(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RA Klein S., Strausberg R.;  
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC044069; AAB44069.1; -  
DR GO: GO:0005869; C:dynactin complex; IEA.  
DR GO: GO:0007017; P:microtubule-based process; IEA.  
DR InterPro: IPR006996; Dynamitin.  
DR Pfam: PF04912; Dynamitin; 1.  
SQ SEQUENCE 338 AA; 38257 MW; 3346E0F0E644B186 CRC64;  
Query Match 84.1%; Score 211; DB 2; Length 338;  
Best Local Similarity 86.0%; Pred. No. 4e-13;  
Matches 43; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
OY 3 KETPOQKRYORLHVEQELTTEVEKIKTTVKESATFEKLTPLVALAKQVAL 52  
DB 31 KETPOQKRYORLHVEQELTTEVEKIKTTVKESATFEKLTPLVALAKQVAL 80  
RESULT 7  
O9PTG6 PRELIMINARY; PRT; 402 AA.  
AC O9PTG6;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
DE Dynamitin.  
GN Name=p50;  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
NCBI\_TaxID=9031;  
OX [1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=20056093; PubMed=10588646;  
RX Valetti C., Metzel D.M., Schrader M., Hasbani M.J., Gill S.R.,  
RA Kreis T.E., Schroer T.A.;  
RT "Role of dynactin in endocytic traffic: effects of dynactin  
RT overexpression and colocalization with CLIP-170.";  
RL Mol. Biol. Cell 10:4107-4120(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Schroer T.A., Gill S.R., Hasbani M.J., Crego C.;  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF200744; AAF13996.1; -  
DR GO: GO:0005869; C:dynactin complex; IEA.

DR GO:0007017; P:microtubule-based process; IEA.  
 DR InterPro: IPR006996; Dyanmitin.  
 DR Pfam: PF04912; Dyanmitin; 1.  
 SQ SEQUENCE 402 AA; 45126 MW; F229C467C630DCB9 CRC64;

Query Match 70.5%; Score 177; DB 2; Length 402;  
 Best Local Similarity 67.3%; Pred. No. 1.2e-09;  
 Matches 35; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

OY 1 GVKETPOQKRYRLHVEOELTTEVERIKITVKESATBEKTLPVLLAKQAL 52  
 DB 96 GAKETPOQKRYRLHVEOELTTEVERIKITVKESATBEKTLPVLLAKQAL 147

## RESULT 8

Q73H1 PRELIMINARY; PRT; 405 AA.

AC Q73H1  
 DT 01-OCT-2003 (TRENBLrel. 25, Created)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TRENBLrel. 25, Last annotation update)  
 DE Similar to dynactin 2 (p50).  
 GN ORFNames=zgc:63867;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
 Brownstein M.J., Usetin T.B., Toshiyuki S., Carninci P., Prange C.,  
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 Fahy U., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,  
 Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,  
 Jones S.J., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
 RA EMBL; BC053120; AAH53120.1;  
 DR EMBL; BC053120; AAH53120.1;  
 DR ZFIN; ZDB-GENE-040426-1279; zgc:63867.  
 DR GO; GO:0005869; C:dynactin complex; IEA.  
 DR GO; GO:0007017; P:microtubule-based process; IEA.  
 DR InterPro: IPR006996; Dyanmitin.  
 DR Pfam; PF04912; Dyanmitin; 1.  
 SQ SEQUENCE 405 AA; 44625 MW; CA00047342500953 CRC64;

Query Match 70.5%; Score 177; DB 2; Length 405;  
 Best Local Similarity 67.3%; Pred. No. 1.2e-09;  
 Matches 35; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

OY 1 GVKETPOQKRYRLHVEOELTTEVERIKITVKESATBEKTLPVLLAKQAL 52  
 DB 96 GAKETPOQKRYRLHVEOELTTEVERIKITVKESATBEKTLPVLLAKQAL 147

## RESULT 9

O6IP53 PRELIMINARY; PRT; 402 AA.

AC O6IP53  
 DT 05-JUL-2004 (TRENBLrel. 27, Created)  
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)  
 DE MG78949 protein.  
 GN Name=MG78949;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
 Brownstein M.J., Usetin T.B., Toshiyuki S., Carninci P., Prange C.,  
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 Fahy U., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,  
 Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,  
 Jones S.J., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 Richardson P.,  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative.";  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 RA EMBL; BC072065; AAH72065.1;  
 DR EMBL; BC072065; AAH72065.1;  
 DR GO; GO:0005874; C:microtubule;  
 DR GO; GO:0000226; P:microtubule cytoskeleton organization and b. . .; IEA.  
 DR InterPro: IPR000435; Tektin.  
 DR Pfam; PF03148; Tektin; 1.  
 DR PRINTS; PR00511; TEKTN.  
 SQ SEQUENCE 402 AA; 46535 MW; F173BC48FC0E798A CRC64;

Query Match 30.3%; Score 76; DB 2; Length 402;  
 Best Local Similarity 46.4%; Pred. No. 16;  
 Matches 13; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

OY 12 RLHVEOELTTEVERIKITVKESATBEK 39  
 DB 336 RLHVEOELTTEVERIKITVKESATBEK 363

RESULT 10  
 O9V4Y9 PRELIMINARY; PRT; 380 AA.  
 ID O9V4Y9

AC Q9V4Y9;  
DT 01-MAY-2000 (TReMBLrel. 13. Created)  
DT 01-MAY-2000 (TReMBLrel. 13. Last sequence update)  
DT 25-OCT-2004 (TReMBLrel. 28. Last annotation update)  
DE CG8269-PA (L007994P).  
GN Name=Dmn; ORFNames=CG8269;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hopkins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazey R.G., Chapple M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
RA Abell J.F., Agbayani A., An H.J., Andrews-Pfankuch C., Baldwin D.,  
RA Ballew R.M., Baas A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.T., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Broctier P.,  
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,  
RA Glodet K., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostali M., Houston K.A., Howland T.J., Mei H., Ibegam C.,  
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,  
RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milhina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paclet J.M.,  
RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puti V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA She B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier B., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=2242665; PubMed=12537568;  
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Change M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hopkins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
RA Paclet J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila  
melanogaster euchromatic genome sequence.";  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celniker S.E.;  
RT "The transposable elements of the Drosophila melanogaster euchromatin:  
a genomic perspective.";  
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).

RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Mirra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochuk S.B.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
systematic review.";  
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX FlyBase;  
RN Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP FlyBase;  
RN Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Berkeley;  
RA Stapleton M., Brockstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
RA Nunoo J., Paclet J., Paragas V., Park S., Phouenavong S., Wan K.,  
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE001835; AAF59034.1; -  
DR EMBL; AY061092; AAL2640.1; -  
DR InFect; Q9V4Y9; -  
DR FlyBase; FBgn0021825; Dmn.  
DR GO; GO:0005869; C:dynactin complex; IEA.  
DR GO; GO:0007017; P:microtubule-based process; IEA.  
DR InterPro; IPR006996; Dynamitin.  
DR Pfam; PF04912; Dynamitin; 1.  
SQ SEQUENCE 360 AA; 41998 MW; CF7E1D3BFF5989C5 CRC64;  
Query Match 28.9%; Score 72.5; DB 2; Length 380;  
Best Local Similarity 48.7%; Pred. No. 34;  
Matches 19; Conservative 5; Mismatches 14; Indels 1; Gaps 1;  
Qy 1 GVKETPOQKQRLHNEVQELTEVEKIKITTKYESATEEK 39  
Db 94 GKEKTPVQKCPQLQIEMNELNEVAALQVD-RKVADSEK 131  
RESULT 11  
ID Q6WVP7 PRELIMINARY; PRT; 800 AA.  
AC Q6WVP7;  
DT 05-JUL-2004 (TReMBLrel. 27. Created)  
DT 05-JUL-2004 (TReMBLrel. 27. Last sequence update)  
DE 05-JUL-2004 (TReMBLrel. 27. Last annotation update)  
DE Probable translation elongation factor EF-G, mitochondrial.  
GN Name=Bl6D18.030;  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Schulte U., Aign V., Hohnsels J., Brandt P., Fartmann B., Holland R.,  
RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX German Neurospora genome project;  
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BX842625; CAB76249.1; -  
DR GO; GO:0005525; F:GTP binding; IEA.

DR GO:0003746; P:translation elongation factor activity; IEA.  
DR GO:0006412; P:protein biosynthesis; IEA.  
DR GO:0006414; P:translational elongation; IEA.  
DR InterPro: IPR004540; EF-G.  
DR InterPro: IPR000640; EFG\_C.  
DR InterPro: IPR009022; EFG\_III\_V.  
DR InterPro: IPR005517; EFG\_IV\_  
DR InterPro: IPR004161; EFTU\_D2.  
DR InterPro: IPR00795; ProCSyn\_GTPbind.  
DR InterPro: IPR005225; Small\_GTP.  
DR InterPro: IPR009000; Translat\_factor.  
DR Pfam: PF00679; EFG\_C; 1.  
DR Pfam: PF03764; EFG\_IV; 1.  
DR Pfam: PF00009; GTP\_EFTU; 1.  
DR Pfam: PF01144; GTP\_EFTU\_D2; 1.  
DR PRINTS: PR00315; ELONGATNFCT.  
DR TIGRFAMs: TIGR00484; EF-G; 1.  
DR TIGRFAMs: TIGR00231; small\_GTP; 1.  
DR PROSITE: PS00301; EFACOR\_GTP; 1.  
KM Elongation factor; GTP-binding; Protein biosynthesis.  
SQ SEQUENCE 800 AA; 88373 MW; F2A7C425F6D4974A CRC64;

Query Match 28.3%; Score 71; DB 2; Length 800;  
Best Local Similarity 26.2%; Pred. No. 1e+02;  
Matches 16; Conservative 14; Mismatches 19; Indels 12; Gaps 1;

QY 2 VKETPOQKRYRLHEVOELTTEV-----EKIKTKVSATBEKLTPLVLAQQL 49  
DB 306 IKELCEQKRGLEIKLADVDEIEMFLBEQTPPEQKKAIRATICKTLPVLMGSAI 365

QY 50 A 50  
DB 366 A 366

RESULT 12  
ID Q7SH14 PRELIMINARY; PRT; 813 AA.  
AC Q7SH14;  
DT 01-MAR-2004 (TREMblrel. 26, Created)  
DT 01-MAR-2004 (TREMblrel. 26, Last sequence update)  
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)  
DE Hypothetical protein.  
GN Name=NCU02955.1;  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=OR74A;  
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,  
RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,  
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,  
RA Qiu D., Ianakiev P., Pedersen D., Braun C.E., Zeller A., Schulte U.,  
RA Selitrenkoff C.P., Kinsey J.A., Braun E.L., Zeller A., Schulte U.,  
RA Koche G.O., Jedd G., Mewes W.J., Staben C., Marcotte E., Greenberg D.,  
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,  
RA Kamal M., Kamvaselis M., Manceau E., Bielke C., Rudd S., Friseman D.,  
RA Krystofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,  
RA Cogoni C., Machino G., Catchside D., Li W., Pratt R.U., Osmann S.A.,  
RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,  
RA Yarden O., Plamann M., Siller S., Dunlap J., Radford A., Armayo R.,  
RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.,  
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.",  
RL Nature 0-0(2003).  
CC -1- CAUTION: The sequence shown here is derived from an  
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL: AABX0100006; EAA36106.1; -.  
DR HSSP: P13551; IFNM.  
DR GO:0005525; F:GTP binding; IEA.

DR GO:0003746; P:translation elongation factor activity; IEA.  
DR GO:0006412; P:protein biosynthesis; IEA.  
DR GO:0006414; P:translational elongation; IEA.  
DR InterPro: IPR004540; EF-G.  
DR InterPro: IPR000640; EFG\_C.  
DR InterPro: IPR009022; EFG\_III\_V.  
DR InterPro: IPR005517; EFG\_IV\_  
DR InterPro: IPR004161; EFTU\_D2.  
DR InterPro: IPR00795; ProCSyn\_GTPbind.  
DR InterPro: IPR005225; Small\_GTP.  
DR InterPro: IPR009000; Translat\_factor.  
DR Pfam: PF00679; EFG\_C; 1.  
DR Pfam: PF03764; EFG\_IV; 1.  
DR Pfam: PF00009; GTP\_EFTU; 1.  
DR Pfam: PF01144; GTP\_EFTU\_D2; 1.  
DR PRINTS: PR00315; ELONGATNFCT.  
DR TIGRFAMs: TIGR00484; EF-G; 1.  
DR TIGRFAMs: TIGR00231; small\_GTP; 1.  
DR PROSITE: PS00301; EFACOR\_GTP; 1.  
KM GTP-binding; Hypothetical protein; Protein biosynthesis.  
SQ SEQUENCE 813 AA; 89810 MW; C9A5093F35B29B05 CRC64;

Query Match 28.3%; Score 71; DB 2; Length 813;  
Best Local Similarity 26.2%; Pred. No. 1e+02;  
Matches 16; Conservative 14; Mismatches 19; Indels 12; Gaps 1;

QY 2 VKETPOQKRYRLHEVOELTTEV-----EKIKTKVSATBEKLTPLVLAQQL 49  
DB 306 IKELCEQKRGLEIKLADVDEIEMFLBEQTPPEQKKAIRATICKTLPVLMGSAI 365

QY 50 A 50  
DB 366 A 366

RESULT 13  
ID Q971V5 PRELIMINARY; PRT; 311 AA.  
AC Q971V5;  
DT 01-DEC-2001 (TREMblrel. 19, Created)  
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)  
DE Hypothetical protein ST1273.  
GN OrderedLocustNames=ST1273;  
OS Sulfolobus tokodaii.  
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
OC Sulfolobus.  
OX NCBI\_TaxID=111955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JCM 10545 / 7;  
RX MEDLINE=21456156; PubMed=11572479;  
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,  
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hoshiyama A., Fukui S.,  
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamuta M., Kato Y.,  
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,  
RA Aoki K.-I., Masuda S., Yanagita M., Nishimura M., Yamagishi A.,  
RA Oshima T., Kikuchi H.;  
RT "Complete genome sequence of an aerobic thermophilic  
RT Crenarchaeon, Sulfolobus tokodaii strain7.",  
RL DNA Res. 8:123-140(2001).  
DR EMBL: AP000985; BAB6315.1; -.  
DR HSSP: P01100; IFOS.  
KM Complete proteome; Hypothetical protein.  
SQ SEQUENCE 311 AA; 37536 MW; 674338A5E031B429 CRC64;

Query Match 27.9%; Score 70; DB 2; Length 311;  
Best Local Similarity 35.3%; Pred. No. 50;  
Matches 18; Conservative 11; Mismatches 18; Indels 4; Gaps 2;

QY 2 VKETPOQKRYRLHEVOELTTEVEKIKTKVSATBEKLTPLVLAQQLAAL 52  
DB 71 IKQLSQKQSLIYVISEIKKEFEQIKNVER---VKETLDPYLQILKTEQL 117



```
RESULT 14
ID 08MOK1 PRELIMINARY; PRT; 751 AA.
AC 08MOK1;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE L008185p.
GN ORFNames=CG10971;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Munzall C.J., Nuno J., Pacleb J., Patagae V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY129439; AAM76181.1; -.
DR FlyBase; FBgn0036309; CG10971.
DR GO; GO:0003779; P:actin binding; IEA.
DR InterPro; IPR002558; ILMEO.
DR Pfam; PF01608; I_LMEO; 1.
DR PROSITE; PS50945; I_LMEO; 1.
SQ SEQUENCE 751 AA; 85362 MW; 7E9BC7661ECC8AA0 CRC64;

Query Match 27.9%; Score 70; DB 2; Length 751;
Best Local Similarity 44.2%; Pred. No. 1.2e+02;
Matches 19; Conservative 9; Mismatches 11; Indels 4; Gaps 2;

Oy 7 OOKYORLHVEOELTTEVEKIKTYKESATSEKLTPTVLAQOL 49
Db 72 KQVNSQLLETKELTNEISKIVNE--KEX-TNLILOKOL 110

RESULT 15
ID 086BS5 PRELIMINARY; PRT; 1087 AA.
AC 086BS5;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE CG10971-PB.
GN ORFNames=CG10971;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN
RP SEQUENCE FROM N.A.
RL MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaber G.L.,
RA Abt'il J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley B.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Broctier P.,
RA Buttle K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
```

```
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
RA Jaisai M., Kallush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reibert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier B., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodruff, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Zhang R.A., Myers R.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195 (2000).
[2]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537558;
RA Ceiniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers R.W., Gibbs R.A., Rubin G.M.,
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RL melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).
[3]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Ceiniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RL a genomic perspective."
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).
[4]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Miara S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Beltenok B.R., Ceiniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RL systematic review."
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083 (2002).
[5]
RN
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003540; AAP49884.1; -.
DR FlyBase; FBgn0036309; CG10971.
DR GO; GO:0003779; P:actin binding; IEA.
DR GO; GO:0005543; F:phospholipid binding; IEA.
DR InterPro; IPR01417; ANTH.
DR InterPro; IPR001026; Epsin_N.
```



DR InterPro; IPR002558; ILMEQ.  
DR InterPro; IPR008943; PI\_bind\_N.  
DR Pfam; PF07651; ANTH; 1.  
DR Pfam; PF01608; I\_LWEO; 1.  
DR ProDom; PD011820; ILMEQ; 1.  
DR ProSite; PS50942; ENTH; 1.  
DR ProSite; PS50945; I\_LWEO; 1.  
SQ SEQUENCE 1087 AA; 124298 MW; C83E709907DFE8E8 CRC64;

Query Match 27.9%; Score 70; DB 2; Length 1087;  
Best Local Similarity 44.2%; Pred. No. 1.8e+02;  
Matches 19; Conservative 9; Mismatches 11; Indels 4; Gaps 2;

Qy 7 QOKXQRLHVEVQELTTEVEKTKTYKESATEEKLTPVLLAKOL 49  
Db 408 KQVNSQLLETPEKELTNEISKIKVNE---KEK-TNLTLOKQI 446

Search completed: November 3, 2005, 22:03:04  
Job time : 94.4866 secs

010001 090000 000000

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: November 3, 2005, 21:51:39 ; Search time 26.6952 Seconds  
(without alignments)  
145,410 Million cell updates/sec

Title: 09782816-3-1-22

Perfect score: 251  
Sequence: 1 GVKETPOOKYQRLHAEVQEL.....ESATEEKLTPVLAQKQALAL 52

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCtus.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66	26.3	142	4	US-09-270-767-43370 Sequence 43370, A
2	64.5	25.7	639	4	US-09-902-540-14908 Sequence 14908, A
3	64	25.5	151	4	US-09-513-999C-7806 Sequence 7806, Ap
4	64	25.5	193	4	US-09-107-532A-5587 Sequence 5587, Ap
5	63	25.1	850	4	US-09-269-858A-8 Sequence 8, Appli
6	63	25.1	851	4	US-09-623-326-43 Sequence 43, Appli
7	62.5	24.9	1270	4	US-09-538-092-1321 Sequence 1321, Ap
8	61.5	24.5	309	4	US-09-248-786A-14974 Sequence 14974, A
9	61	24.3	442	3	US-09-081-689-2 Sequence 2, Appli
10	61	24.3	442	3	US-09-305-984-14 Sequence 14, Appli
11	61	24.3	442	3	US-09-073-541A-14 Sequence 14, Appli
12	61	24.3	442	4	US-09-493-940-14 Sequence 14, Appli
13	61	24.3	442	4	US-09-583-110-5317 Sequence 11, Appli
14	61	24.3	446	4	US-09-107-433-3720 Sequence 3720, Ap
15	61	24.3	644	1	US-08-487-890A-6 Sequence 6, Appli
16	61	24.3	644	2	US-08-478-435-6 Sequence 6, Appli
17	61	24.3	644	2	US-08-337-483-6 Sequence 6, Appli
18	61	24.3	644	2	US-08-478-373-6 Sequence 6, Appli
19	61	24.3	644	3	US-08-474-671-6 Sequence 6, Appli
20	61	24.3	644	3	US-08-483-577A-6 Sequence 6, Appli
21	61	24.3	644	3	US-08-897-438-6 Sequence 6, Appli
22	61	24.3	644	3	US-08-637-654-6 Sequence 6, Appli
23	61	24.3	644	3	US-08-649-518-6 Sequence 6, Appli
24	60.5	24.1	87	3	US-09-284-033-3 Sequence 3, Appli
25	60.5	24.1	87	3	US-08-729-834B-3 Sequence 3, Appli
26	60.5	24.1	1217	4	US-09-949-016-7454 Sequence 7454, Ap
27	59.5	23.7	424	4	US-09-286-981B-14 Sequence 14, Appli

28	59.5	23.7	425	4	US-09-286-981B-13 Sequence 13, Appli
29	59.5	23.7	426	4	US-09-286-981B-12 Sequence 12, Appli
30	59.5	23.7	721	4	US-09-248-796A-18800 Sequence 18800, A
31	59	23.5	451	4	US-09-270-767-45139 Sequence 45139, A
32	59	23.5	2101	1	US-08-466-390-4 Sequence 4, Appli
33	59	23.5	2101	1	US-08-470-950-4 Sequence 4, Appli
34	59	23.5	2101	1	US-08-467-781-4 Sequence 4, Appli
35	59	23.5	2101	1	US-08-195-487-4 Sequence 4, Appli
36	59	23.5	2101	2	US-08-483-924-4 Sequence 4, Appli
37	59	23.5	2101	3	US-09-452-294-1 Sequence 1, Appli
38	59	23.5	2101	3	PCT-US93-06160-4 Sequence 4, Appli
39	59	23.5	2107	4	US-09-949-016-7646 Sequence 7646, Ap
40	59	23.5	2107	4	US-09-284-033-2 Sequence 2, Appli
41	58.5	23.3	399	3	US-08-729-834B-2 Sequence 2, Appli
42	58.5	23.3	399	3	US-09-949-016-6466 Sequence 6466, Ap
43	58.5	23.3	404	4	US-09-949-016-8325 Sequence 8325, Ap
44	58.5	23.3	487	3	US-09-724-224-8 Sequence 8, Appli
45	58	23.1			

## ALIGNMENTS

```
RESULT 1
US-09-270-767-43370
Sequence 43370, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 43370
LENGTH: 142
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-43370

Query Match      26.3%; Score 66; DB 4; Length 142;
Best Local Similarity 50.0%; Pred. No. 1.6;
Matches 16; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

Cy      1 GVKETPOOKYQRLHAEVQELTTEVEKIKTKTK 32
Db      111 GKEKTPQOKQRLHAEVQELTTEVEKIKTKTK 142

RESULT 2
US-09-902-540-14908
Sequence 14908, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US/09/902,540
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 14908
LENGTH: 639
TYPE: PRT
ORGANISM: Myxococcus xanthus
US-09-902-540-14908

Query Match      25.7%; Score 64.5; DB 4; Length 639;
```

Best Local Similarity 31.9%; Pred. No. 14;  
Matches 15; Conservative 11; Mismatches 12; Indels 9; Gaps 1;

Qy 1 GVKETPOQKYORLHEVOELTTEVEK-----IKTVKESATEE 38  
Db 151 GALEBAVEOYQOALADVERLTSELEAASAEKDSIGLRTAOLEAALEE 197

## RESULT 3

US-09-513-999C-7806  
; Sequence 7806, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 7806  
; LENGTH: 151  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: -25...-1  
; OTHER INFORMATION: score 3.9  
; OTHER INFORMATION: seq RVGLLLGGGAVYG/SR  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 94  
; OTHER INFORMATION: Xaa=Ile or Lys or Asn or Arg or Ser or Thr  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 103  
; OTHER INFORMATION: Xaa=Glu or Lys  
US-09-513-999C-7806

Query Match 25.5%; Score 64; DB 4; Length 151;  
Best Local Similarity 40.5%; Pred. No. 3;  
Matches 17; Conservative 7; Mismatches 8; Indels 10; Gaps 1;

Qy 9 KYORLHEVOELTTEVEKIKTVKESA-----TEEKL 40  
Db 86 KYKRLKAEVKESKLEKKKETITSESARQOKKXIERQBEKL 127

## RESULT 4

US-09-107-532A-5587  
; Sequence 5587, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7110  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD/ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 5587:

SEQUENCE CHARACTERISTICS:  
LENGTH: 193 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...193  
SEQUENCE DESCRIPTION: SEQ ID NO: 5587:

US-09-107-532A-5587  
Query Match 25.5%; Score 64; DB 4; Length 193;  
Best Local Similarity 40.5%; Pred. No. 3.9;  
Matches 15; Conservative 6; Mismatches 14; Indels 2; Gaps 1;

Qy 2 VKETPOQKYORLHEVOELTTEVEKIKTVKESATEE 38  
Db 70 VSEIVEQVEIEELIE--EPTAIEIVNVETILIEKTEE 104

## RESULT 5

US-09-269-858A-8  
; Sequence 8, Application US/09269858A  
; Patent No. 6692932  
; GENERAL INFORMATION:  
; APPLICANT: Angerer, Bernhard  
; APPLICANT: Ankenbauer, Waltraud  
; APPLICANT: Boehringer, M.  
; APPLICANT: Bonch-Osmolovskaya, Elizaveta  
; APPLICANT: Markau, Ursula  
; APPLICANT: Reiser, Astrid  
; APPLICANT: Schmitz-Agnefian, Gudrun  
; APPLICANT: Svetlichny, Vitaly  
; TITLE OF INVENTION: Thermostable DNA Polymerase From Anaerocellum Thermophilum  
; FILE REFERENCE: 4452  
; CURRENT APPLICATION NUMBER: US/09/269,858A  
; CURRENT FILING DATE: 1999-06-10  
; PRIOR APPLICATION NUMBER: PCT/EP97/05390  
; PRIOR FILING DATE: 1996-10-03  
; PRIOR APPLICATION NUMBER: EP/96115877  
; PRIOR FILING DATE: 1996-10-03  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 8  
; LENGTH: 850  
; TYPE: PRT  
; ORGANISM: Abecus herberti  
US-09-269-858A-8

Query Match 25.1%; Score 63; DB 4; Length 850;  
Best Local Similarity 30.9%; Pred. No. 28;  
Matches 17; Conservative 15; Mismatches 19; Indels 4; Gaps 2;

OY 1 GVKETPOOKYORLHE---VOELTTEVEKITVKSATBEKLPVLLAKOLAL 52  
Db 192 GVKIGIKSKAKLBEYSLEIYQNDKIKSIREKLEAGK-DMAFLSKRLATI 245

RESULT 6  
US-09-623-326-43  
; Sequence 43, Application US/09623326  
; Patent No. 6607883  
; GENERAL INFORMATION:  
; APPLICANT: Frey et al.  
; TITLE OF INVENTION: Polymerase Chimerae  
; FILE REFERENCE: 4894  
; CURRENT APPLICATION NUMBER: US/09/623,326  
; PRIOR FILING DATE: 2001-02-08  
; PRIOR APPLICATION NUMBER: DE 198 10 879.6  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 43  
; LENGTH: 851  
; TYPE: PRT  
; ORGANISM: Anaerocellum thermophilum  
US-09-623-326-43

Query Match 25.1%; Score 63; DB 4; Length 851;  
Best Local Similarity 30.9%; Pred. No. 28;  
Matches 17; Conservative 15; Mismatches 19; Indels 4; Gaps 2;

OY 1 GVKETPOOKYORLHE---VOELTTEVEKITVKSATBEKLPVLLAKOLAL 52  
Db 192 GVKIGIKSKAKLBEYSLEIYQNDKIKSIREKLEAGK-DMAFLSKRLATI 245

RESULT 7  
US-09-538-092-1321  
; Sequence 1321, Application US/09538092  
; Patent No. 6753314  
; GENERAL INFORMATION:  
; APPLICANT: Mansfield, Traci A.  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538,092  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/178,965  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CurataseqFormatter Version 0.9  
; SEQ ID NO 1321  
; LENGTH: 1270  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (0)..(0)  
; OTHER INFORMATION: Polypeptide Accession Number Q14203  
US-09-538-092-1321

Query Match 24.9%; Score 62.5; DB 4; Length 1270;  
Best Local Similarity 34.8%; Pred. No. 52;  
Matches 16; Conservative 11; Mismatches 18; Indels 1; Gaps 1;

OY 4 ETPQOKYORLHEVOELTTEVEKITVKSATBEKLPVLLAKOL 49  
Db 311 ESLQOEVALKERVDELITLILKATIEKSGDAASSTYOL-KOL 355

RESULT 8  
US-09-248-796A-14974  
; Sequence 14974, Application US/09248796A

; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; PRIOR FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 14974  
; LENGTH: 309  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-14974

Query Match 24.5%; Score 61.5; DB 4; Length 309;  
Best Local Similarity 41.2%; Pred. No. 13;  
Matches 14; Conservative 8; Mismatches 9; Indels 3; Gaps 1;

OY 3 KETPOO---KYORLHEVOELTTEVEKITVKE 33  
Db 271 EETPIQDNPXRILHITDLTSDIEDIEPTPE 304

RESULT 9  
US-09-081-689-2  
; Sequence 2, Application US/09081689  
; Patent No. 6165992  
; GENERAL INFORMATION:  
; APPLICANT: Wallis, Nicola G.  
; APPLICANT: Zalacain, Magdalena  
; APPLICANT: Throup, John  
; APPLICANT: Biswas, Sanjoy  
; TITLE OF INVENTION: Histidine Kinase  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechert, Price & Rhoads  
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103-2793  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/081,689  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/048,347  
; FILING DATE: 30-MAY-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Falk, Stephen T  
; REGISTRATION NUMBER: 36,795  
; REFERENCE/DOCKET NUMBER: GM10009  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-994-2488  
; TELEFAX: 215-994-2222  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 442 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-081-689-2

```
Query Match      24.3%; Score 61; DB 3; Length 442;
Best Local Similarity 33.3%; Pred. No. 23;
Matches 18; Conservative 6; Mismatches 24; Indels 6; Gaps 1;

Oy      2 VKETPOQKYORL-----LHEVOELTTEVEKIKTTVESATEEKLTPVLLAKOL 49
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db      200 LKEQINSLYOHLTVIADLHEKNEAIIQLERKVFELRGASHELKTPLASIKIL 253

RESULT 10
US-09-305-984-14
; Sequence 14, Application US/09305984B
; Patent No. 6331407
; GENERAL INFORMATION:
; APPLICANT: No. 6331407ak, Rodger
; APPLICANT: Toumanen, Elaine
; TITLE OF INVENTION: NOVEL ANTIBIOTICS AND METHODS OF USING THE SAME
; FILE REFERENCE: 1340-1-016N1
; CURRENT FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: 60/084,399
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: 09/305,984
; EARLIER FILING DATE: 1999-05-05
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 14
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-305-984-14

Query Match      24.3%; Score 61; DB 3; Length 442;
Best Local Similarity 33.3%; Pred. No. 23;
Matches 18; Conservative 6; Mismatches 24; Indels 6; Gaps 1;

Oy      2 VKETPOQKYORL-----LHEVOELTTEVEKIKTTVESATEEKLTPVLLAKOL 49
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db      200 LKEQINSLYOHLTVIADLHEKNEAIIQLERKVFELRGASHELKTPLASIKIL 253

RESULT 11
US-09-073-541A-14
; Sequence 14, Application US/09073541A
; Patent No. 6448224
; GENERAL INFORMATION:
; APPLICANT: No. 6448224ak, Rodger
; APPLICANT: Toumanen, Elaine
; TITLE OF INVENTION: NOVEL ANTIBIOTICS AND METHODS OF USING THE SAME
; FILE REFERENCE: 1340-1-016
; CURRENT APPLICATION NUMBER: US/09/073,541A
; CURRENT FILING DATE: 1998-05-06
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 14
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-073-541A-14

Query Match      24.3%; Score 61; DB 4; Length 442;
Best Local Similarity 33.3%; Pred. No. 23;
Matches 18; Conservative 6; Mismatches 24; Indels 6; Gaps 1;

Oy      2 VKETPOQKYORL-----LHEVOELTTEVEKIKTTVESATEEKLTPVLLAKOL 49
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db      200 LKEQINSLYOHLTVIADLHEKNEAIIQLERKVFELRGASHELKTPLASIKIL 253

RESULT 12
US-09-493-940-14
; Sequence 14, Application US/09493940
```

```
; Patent No. 6630583
; GENERAL INFORMATION:
; APPLICANT: No. 6630583ak, Rodger
; APPLICANT: Toumanen, Elaine
; TITLE OF INVENTION: NOVEL ANTIBIOTICS AND METHODS OF USING THE SAME
; FILE REFERENCE: 1340-1-016N1
; CURRENT APPLICATION NUMBER: US/09/493,940
; CURRENT FILING DATE: 2000-01-28
; EARLIER APPLICATION NUMBER: 60/084,399
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: 09/305,984
; EARLIER FILING DATE: 1999-05-05
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 14
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-493-940-14

Query Match      24.3%; Score 61; DB 4; Length 442;
Best Local Similarity 33.3%; Pred. No. 23;
Matches 18; Conservative 6; Mismatches 24; Indels 6; Gaps 1;

Oy      2 VKETPOQKYORL-----LHEVOELTTEVEKIKTTVESATEEKLTPVLLAKOL 49
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db      200 LKEQINSLYOHLTVIADLHEKNEAIIQLERKVFELRGASHELKTPLASIKIL 253

RESULT 13
US-09-583-110-5317
; Sequence 5317, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 5317
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-5317

Query Match      24.3%; Score 61; DB 4; Length 442;
Best Local Similarity 33.3%; Pred. No. 23;
Matches 18; Conservative 6; Mismatches 24; Indels 6; Gaps 1;

Oy      2 VKETPOQKYORL-----LHEVOELTTEVEKIKTTVESATEEKLTPVLLAKOL 49
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db      200 LKEQINSLYOHLTVIADLHEKNEAIIQLERKVFELRGASHELKTPLASIKIL 253

RESULT 14
US-09-107-433-3720
; Sequence 3720, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNO
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: <Unknown>  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: <Unknown>  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,433  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085131  
FILING DATE: May 12, 1998  
APPLICATION NUMBER: 60/051553  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Arinello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-011  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 3720:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 446 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pneumoniae  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (B) LOCATION 1...446  
SEQUENCE DESCRIPTION: SEQ ID NO: 3720:  
US-09-107-433-3720

Query March 24.3% Score 61; DB 4; Length 446;  
Best Local Similarity 33.3%; Pred. NO. 23;  
Matches 18; Conservative 6; Mismatches 24; Indels 6; Gaps 1;

QY 2 VKETPOOKYQRL-----LHEVOELTEVERIKITTVKESATEBKLTPLYLAKOL 49  
DB 204 LKEQINSYQHLTVIADLHEKNEAILQLKQKVFPLRGASHELKTPLASIKIL 257

RESULT 15  
US-08-487-890A-6  
Sequence 6, Application US/08487890A  
Patent No. 5708149  
GENERAL INFORMATION:  
APPLICANT: Loosmore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Mirdin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,890A  
FILING DATE: 07-JUN-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/175,116  
FILING DATE: 29-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,968  
FILING DATE: 08-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-466 MIS:jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 644 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-487-890A-6

Query March 24.3% Score 61; DB 1; Length 644;  
Best Local Similarity 32.6%; Pred. NO. 36;  
Matches 15; Conservative 8; Mismatches 17; Indels 6; Gaps 1;

DD 2 VKETPOOKYQRLHEVOELTEVE-----KIKTVKESATEBKLT 41  
244 LREIDANKSQNRTHKLYDLEADVHNSNPRGKVKTKKSSSEHPT 289

Search completed: November 3, 2005, 22:05:56  
Job time : 27.6952 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2005, 21:53:15 ; Search time 93.7112 Seconds  
(without alignments)  
232.174 Million cell updates/sec

Title: 09782816-3-1-22

Perfect score: 251  
Sequence: 1 GVKEETPOOKYORLHVEVQEL.....ESATEBKLPVLAKQALAL 52

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09A\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	251	100.0	401	16 US-10-408-765A-1369	Sequence 1369, App
2	251	100.0	406	10 US-09-782-816A-53	Sequence 53, Appl
3	251	100.0	465	9 US-09-925-298-676	Sequence 676, App
4	251	100.0	465	14 US-10-102-806-676	Sequence 676, App
5	248	98.8	183	10 US-09-782-816A-54	Sequence 54, Appl
6	248	98.8	224	16 US-10-425-115-115831	Sequence 51, Appl
7	246	98.0	52	10 US-09-782-816A-51	Sequence 51, Appl
8	231	92.0	134	14 US-10-106-698-6730	Sequence 6730, Appl
9	113	45.0	22	10 US-09-782-816A-3	Sequence 3, Appl
10	107	42.6	21	10 US-09-782-816A-4	Sequence 4, Appl
11	103	41.0	20	10 US-09-782-816A-5	Sequence 5, Appl

	12	103	41.0	23	10	US-09-782-816A-1	Sequence 1, Appl
	13	98	39.0	19	10	US-09-782-816A-6	Sequence 6, Appl
	14	93	37.1	18	10	US-09-782-816A-7	Sequence 7, Appl
	15	88	35.1	17	10	US-09-782-816A-8	Sequence 8, Appl
	16	81	32.3	16	10	US-09-782-816A-9	Sequence 9, Appl
	17	76	30.3	15	10	US-09-782-816A-10	Sequence 10, Appl
	18	72.5	28.9	53	10	US-09-782-816A-52	Sequence 52, Appl
	19	72.5	28.9	380	10	US-09-782-816A-56	Sequence 56, Appl
	20	72.5	28.9	380	20	US-11-097-143-4056	Sequence 4056, App
	21	71	28.3	14	10	US-09-782-816A-11	Sequence 11, Appl
	22	70.5	28.1	348	16	US-10-425-115-210465	Sequence 210465, A
	23	70.5	28.1	356	15	US-10-425-114-63356	Sequence 63356, A
	24	70	27.9	1087	20	US-11-097-143-22866	Sequence 22866, A
	25	70	27.9	1087	20	US-11-097-143-22869	Sequence 22869, A
	26	68	27.1	2368	9	US-09-815-242-5635	Sequence 5635, App
	27	68	27.1	2368	9	US-09-815-242-12389	Sequence 12389, A
	28	67	26.7	2478	9	US-09-815-242-5816	Sequence 5816, App
	29	67	26.7	2478	9	US-09-815-242-12967	Sequence 12967, A
	30	67	26.7	2478	17	US-10-470-0488-220	Sequence 220, App
	31	67	26.7	2481	15	US-10-282-122A-43762	Sequence 43762, A
	32	66	26.3	13	10	US-09-782-816A-12	Sequence 12, Appl
	33	66	26.3	386	15	US-10-282-122A-71967	Sequence 71967, A
	34	65.5	26.1	576	15	US-10-424-599-274041	Sequence 274041, A
	35	65.5	26.1	1837	15	US-10-369-493-22734	Sequence 22734, A
	36	64.5	25.7	348	16	US-10-437-963-139713	Sequence 139713, A
	37	64	25.5	185	16	US-10-842-740-3	Sequence 3, Appl
	38	64	25.5	188	10	US-09-284-320-12	Sequence 12, Appl
	39	64	25.5	188	14	US-10-028-072-2	Sequence 2, Appl
	40	64	25.5	188	14	US-10-140-808-2	Sequence 2, Appl
	41	64	25.5	188	14	US-10-121-049-2	Sequence 2, Appl
	42	64	25.5	188	14	US-10-123-904-2	Sequence 2, Appl
	43	64	25.5	188	14	US-10-140-470-2	Sequence 2, Appl
	44	64	25.5	188	14	US-10-175-746-2	Sequence 2, Appl
	45	64	25.5	188	14	US-10-176-918-2	Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-10-408-765A-1369  
; Sequence 1369, Application US/10408765A  
; Publication No. US20040101874A1  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Fahy, Boim D.  
; APPLICANT: Zhang, Bing  
; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.  
; APPLICANT: Glenn, Gary M.  
; APPLICANT: Martnock, Dale B.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
; FILE REFERENCE: 660088.465  
; CURRENT APPLICATION NUMBER: US/10/408,765A  
; CURRENT FILING DATE: 2003-04-04  
; NUMBER OF SEQ ID NOS: 3077  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1369  
; LENGTH: 401  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-408-765A-1369

Query Match 100.0%; Score 251; DB 16; Length 401;  
Beat local similarity 100.0%; Pred. No. 1.5e-19;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 GVKEETPOOKYORLHVEVQELTTEVEKIKTKYKESATEBKLPVLAKQALAL 52  
DB 94 GVKEETPOOKYORLHVEVQELTTEVEKIKTKYKESATEBKLPVLAKQALAL 145

RESULT 2  
US-09-782-816A-53  
; Sequence 53, Application US/09782816A  
; Publication No. US20030032771A1  
; GENERAL INFORMATION:  
; APPLICANT: Sharp, David J.  
; APPLICANT: Rogers, Gregory C.  
; APPLICANT: Scholey, Jonathan M.  
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR  
; TITLE OF INVENTION: PROLIFERATION  
; FILE REFERENCE: U0069.001A  
; CURRENT APPLICATION NUMBER: US/09/782.816A  
; CURRENT FILING DATE: 2001-02-14  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 53  
; LENGTH: 406  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-782-816A-53

Query Match 100.0%; Score 251; DB 10; Length 406;  
Best Local Similarity 100.0%; Pred. No. 1.6e-19;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVKETPOOKYORLHVEOELTTEVEKIKTTVKSATSEKLTPLVLAQALAL 52  
Db 99 GVKETPOOKYORLHVEOELTTEVEKIKTTVKSATSEKLTPLVLAQALAL 150

RESULT 3  
US-09-925-298-676  
; Sequence 676, Application US/09925298  
; Publication No. US20020039764A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA103  
; CURRENT APPLICATION NUMBER: US/09/925.298  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05881  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 846  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 676  
; LENGTH: 465  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (5)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (6)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (16)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-298-676

Query Match 100.0%; Score 251; DB 9; Length 465;  
Best Local Similarity 100.0%; Pred. No. 1.8e-19;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVKETPOOKYORLHVEOELTTEVEKIKTTVKSATSEKLTPLVLAQALAL 52  
Db 158 GVKETPOOKYORLHVEOELTTEVEKIKTTVKSATSEKLTPLVLAQALAL 209

RESULT 4  
US-10-102-806-676

; Sequence 676, Application US/10102806  
; Publication No. US20030054421A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA103P1C1  
; CURRENT APPLICATION NUMBER: US/10/102.806  
; CURRENT FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: 09/925,298  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05881  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 846  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 676  
; LENGTH: 465  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (5)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (6)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (16)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-102-806-676

Query Match 100.0%; Score 251; DB 14; Length 465;  
Best Local Similarity 100.0%; Pred. No. 1.8e-19;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVKETPOOKYORLHVEOELTTEVEKIKTTVKSATSEKLTPLVLAQALAL 52  
Db 158 GVKETPOOKYORLHVEOELTTEVEKIKTTVKSATSEKLTPLVLAQALAL 209

RESULT 5  
US-09-782-816A-54  
; Sequence 54, Application US/09782816A  
; Publication No. US20030032771A1  
; GENERAL INFORMATION:  
; APPLICANT: Sharp, David J.  
; APPLICANT: Rogers, Gregory C.  
; APPLICANT: Scholey, Jonathan M.  
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR  
; TITLE OF INVENTION: PROLIFERATION  
; FILE REFERENCE: U0069.001A  
; CURRENT APPLICATION NUMBER: US/09/782.816A  
; CURRENT FILING DATE: 2001-02-14  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 54  
; LENGTH: 183  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-782-816A-54

Query Match 98.8%; Score 248; DB 10; Length 183;  
Best Local Similarity 98.1%; Pred. No. 1.3e-19;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVKETPOOKYORLHVEOELTTEVEKIKTTVKSATSEKLTPLVLAQALAL 52  
Db 94 GVKETPOOKYORLHVEOELTTEVEKIKTTVKSATSEKLTPLVLAQALAL 145

RESULT 6  
US-10-425-115-315831

; Sequence 315831, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Placids  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425.115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 315831  
; LENGTH: 224  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_51109C.1.pcp  
US-10-425-115-315831

Query Match 98.8%; Score 248; DB 16; Length 224;  
Best Local Similarity 98.1%; Pred. No. 1.7e-19;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GVKETPOOKYORLHVEOELTTEVEKIKTTVKESATEEKLTPVLIAKOLAL 52  
Db 47 GVKETPOOKYORLHVEOELTTEVEKIKTTVKESATEEKLTPVLIAKOLAL 98

## RESULT 7

US-09-782-816A-51  
; Sequence 51, Application US/09782816A  
; Publication No. US20030032771A1  
; GENERAL INFORMATION:  
; APPLICANT: Sharp, David J.  
; APPLICANT: Rogers, Gregory C.  
; APPLICANT: Scholey, Jonathan M.  
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR  
; FILE REFERENCE: UC069.001A  
; CURRENT APPLICATION NUMBER: US/09/782.816A  
; CURRENT FILING DATE: 2001-02-14  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 51  
; LENGTH: 52  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 44  
; OTHER INFORMATION: Xaa = Val or Leu  
; OTHER INFORMATION: A sequence conserved among Homo sapiens and Mus  
; OTHER INFORMATION: musculus.  
US-09-782-816A-51

Query Match 98.0%; Score 246; DB 10; Length 52;  
Best Local Similarity 98.1%; Pred. No. 5e-20;  
Matches 51; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GVKETPOOKYORLHVEOELTTEVEKIKTTVKESATEEKLTPVLIAKOLAL 52  
Db 1 GVKETPOOKYORLHVEOELTTEVEKIKTTVKESATEEKLTPVLIAKOLAL 52

## RESULT 8

US-10-106-698-6730  
; Sequence 6730, Application US/10106698  
; Publication No. US20030109690A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide

; FILE REFERENCE: PA005P1  
; CURRENT APPLICATION NUMBER: US/10/106.698  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: PCT/US00/26524  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/157,137  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: US 60/163,280  
; PRIOR FILING DATE: 1999-11-03  
; NUMBER OF SEQ ID NOS: 8564  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 6730  
; LENGTH: 134  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (126)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-106-698-6730

Query Match 92.0%; Score 231; DB 14; Length 134;  
Best Local Similarity 100.0%; Pred. No. 7.1e-18;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5 TPQOKYORLHVEOELTTEVEKIKTTVKESATEEKLTPVLIAKOLAL 52  
Db 1 TPQOKYORLHVEOELTTEVEKIKTTVKESATEEKLTPVLIAKOLAL 48

## RESULT 9

US-09-782-816A-3  
; Sequence 3, Application US/09782816A  
; Publication No. US20030032771A1  
; GENERAL INFORMATION:  
; APPLICANT: Sharp, David J.  
; APPLICANT: Rogers, Gregory C.  
; APPLICANT: Scholey, Jonathan M.  
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR  
; FILE REFERENCE: UC069.001A  
; CURRENT APPLICATION NUMBER: US/09/782.816A  
; CURRENT FILING DATE: 2001-02-14  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 22  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: A sequence conserved among Homo sapiens and Mus  
; OTHER INFORMATION: musculus.  
US-09-782-816A-3

Query Match 45.0%; Score 113; DB 10; Length 22;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GVKETPOOKYORLHVEOELTTEVEKIKTTVKESATEEKLTPVLIAKOLAL 22  
Db 1 GVKETPOOKYORLHVEOELTTEVEKIKTTVKESATEEKLTPVLIAKOLAL 22

## RESULT 10

US-09-782-816A-4  
; Sequence 4, Application US/09782816A  
; Publication No. US20030032771A1  
; GENERAL INFORMATION:  
; APPLICANT: Sharp, David J.  
; APPLICANT: Rogers, Gregory C.  
; APPLICANT: Scholey, Jonathan M.  
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR

```
; FILE REFERENCE: UC069.001A
; CURRENT APPLICATION NUMBER: US/09/782,816A
; CURRENT FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: A sequence conserved among Homo sapiens and Mus
; OTHER INFORMATION: musculus.
US-09-782-816A-4
```

```
Query Match          42.6%; Score 107; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 4,7e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      2 VKETPOQKYQRLHVEVOELTT 22
         |||||
Db      1 VKETPOQKYQRLHVEVOELTT 21
```

```
RESULT 11
US-09-782-816A-5
; Sequence 5, Application US/09782816A
; Publication No. US20030032771A1
; GENERAL INFORMATION:
; APPLICANT: Sharp, David J.
; APPLICANT: Rogers, Gregory C.
; APPLICANT: Scholey, Jonathan M.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
; TITLE OF INVENTION: PROLIFERATION
; FILE REFERENCE: UC069.001A
; CURRENT APPLICATION NUMBER: US/09/782,816A
; CURRENT FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: A sequence conserved among Homo sapiens and Mus
; OTHER INFORMATION: musculus.
US-09-782-816A-5
```

```
Query Match          41.0%; Score 103; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      3 KETPOQKYQRLHVEVOELTT 22
         |||||
Db      1 KETPOQKYQRLHVEVOELTT 20
```

```
RESULT 12
US-09-782-816A-1
; Sequence 1, Application US/09782816A
; Publication No. US20030032771A1
; GENERAL INFORMATION:
; APPLICANT: Sharp, David J.
; APPLICANT: Rogers, Gregory C.
; APPLICANT: Scholey, Jonathan M.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
; TITLE OF INVENTION: PROLIFERATION
; FILE REFERENCE: UC069.001A
; CURRENT APPLICATION NUMBER: US/09/782,816A
; CURRENT FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 23
```

```
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 22
; OTHER INFORMATION: Xaa = Val or Leu
; OTHER INFORMATION: The sequence is a Homo sapiens sequence when Xaa
; OTHER INFORMATION: represents Leu and a Mus musculus sequence when
; OTHER INFORMATION: Xaa represents Val.
US-09-782-816A-1
```

```
Query Match          41.0%; Score 103; DB 10; Length 23;
Best Local Similarity 95.7%; Pred. No. 0.00015;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy      23 EVKIKTTVESATBEKLTVPYL 45
         |||||
Db      1 EVKIKTTVESATBEKLTVPYL 23
```

```
RESULT 13
US-09-782-816A-6
; Sequence 6, Application US/09782816A
; Publication No. US20030032771A1
; GENERAL INFORMATION:
; APPLICANT: Sharp, David J.
; APPLICANT: Rogers, Gregory C.
; APPLICANT: Scholey, Jonathan M.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
; TITLE OF INVENTION: PROLIFERATION
; FILE REFERENCE: UC069.001A
; CURRENT APPLICATION NUMBER: US/09/782,816A
; CURRENT FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: A sequence conserved among Homo sapiens and Mus
; OTHER INFORMATION: musculus.
US-09-782-816A-6
```

```
Query Match          39.0%; Score 98; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      4 ETPOQKYQRLHVEVOELTT 22
         |||||
Db      1 ETPOQKYQRLHVEVOELTT 19
```

```
RESULT 14
US-09-782-816A-7
; Sequence 7, Application US/09782816A
; Publication No. US20030032771A1
; GENERAL INFORMATION:
; APPLICANT: Sharp, David J.
; APPLICANT: Rogers, Gregory C.
; APPLICANT: Scholey, Jonathan M.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
; TITLE OF INVENTION: PROLIFERATION
; FILE REFERENCE: UC069.001A
; CURRENT APPLICATION NUMBER: US/09/782,816A
; CURRENT FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
```

OTHER INFORMATION: A sequence conserved among Homo sapiens and Mus  
OTHER INFORMATION: musculus.  
US-09-782-816A-7

Query Match 37.1%; Score 93; DB 10; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TPQKRYRLHVEVQLTT 22  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 TPQKRYRLHVEVQLTT 18

RESULT 15  
US-09-782-816A-8  
Sequence 8, Application US/09782816A  
Publication No. US20030032771A1  
GENERAL INFORMATION:  
APPLICANT: Sharp, David J.  
APPLICANT: Rogers, Gregory C.  
APPLICANT: Scholley, Jonathan M.  
TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR  
PROLIFERATION  
FILE REFERENCE: UC069,001A  
CURRENT APPLICATION NUMBER: US/09/782,816A  
CURRENT FILING DATE: 2001-02-14  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 17  
TYPE: PRT  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: A sequence conserved among Homo sapiens and Mus  
US-09-782-816A-8

Query Match 35.1%; Score 88; DB 10; Length 17;  
Best Local Similarity 100.0%; Pred. No. 0.0048;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PQQKRYRLHVEVQLTT 22  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 PQQKRYRLHVEVQLTT 17

Search completed: November 3, 2005, 22:11:40  
Job time : 94.7112 secs

**This Page Blank (uspto)**